

When to cherish white elephants

Nuclear power stations in several places have become the most conspicuous of unused resources. What will happen to them if it emerges that the greenhouse effect is a reality?

HARDLY anybody would have known of the Long Island Lighting Company (called Lilco) if it were not the owner of the Shoreham nuclear plant — ready to operate for the past three years, never used and now likely to be abandoned. That appears to be the objective of the protracted negotiations between Lilco and the Long Island Power Board, a regional utility operated by the state of New York. On one version of the talks, the power board will simply take over Lilco (but not Shoreham). On another, Lilco will agree not to operate the power plant if it can secure supplies of electricity from elsewhere (and increase its charges to its customers in the process). While the Shoreham plant appears to meet the technical safety standards required of new nuclear installations in the United States, it appears not to be possible for its owners to satisfy state and federal requirements (made more stringent after Three Mile Island, when Shoreham was almost complete) for the evacuation of the neighbouring population in the case of a nuclear accident. Meanwhile, the company is to face a trial next September of a civil lawsuit alleging that it defrauded the regulatory authorities and its customers in the building of the Shoreham plant. It is just over a month since the Seabrook plant in New Hampshire was also abandoned, but for different reasons.

Profligacy on such a scale is even these days uncommon, even though the US nuclear industry has an unenviable record of ill-judged and mismanaged investment in nuclear plants. But until recently, the government of Austria seemed to hold the record for deciding not to use a nuclear plant whose technical safety was not disputed. (Now, after Chernobyl, that plant is being dismantled.) In Sweden, where it has been decided to build no further nuclear plants, the existing plant is being operated successfully and, unsurprisingly, safely, which is a sensible use to make of expensive capital equipment. On Long Island, the cost of Shoreham is put at \$5,000 million (allowing for interest on the construction cost). That is a lot of anybody's money to discard. The present value of the plant, the difference between income and operating costs amortized to the present, is probably several times as much, no doubt enough to provide each of the families in the disputed evacuation zone with a new automobile and new highways on which to drive them.

Under the theoretically hands-off system of government of the United States, the loss of the cost of Shoreham will be borne, in the first instance, by private persons and companies — the shareholders of Lilco and those who bought the bonds whose proceeds were used to build the plant. Ultimately, however, the costs will be more widely spread. To the extent that those who buy bonds as investments will be persuaded by Shoreham and its predecessors that utilities operating nuclear plants are an unpredictable risk, the cost of some conventional power developments will be increased while that of financing nuclear developments of any kind may be made uneconomically high. That will not matter if nuclear power, in the circumstances of the United States, is uneconomic now and likely to be so for the rest of time. But what if circumstances should change? What, for example, if the effect of accumulating carbon dioxide in the atmosphere is indeed shown to cause climatic change?

The many who complain that it is facetious to suggest even

indirectly that the remote risk of a nuclear accident can be made tolerable on economic grounds may find themselves differently impelled if it becomes clear that the only economically feasible way of mitigating the damage that might be done by carbon dioxide is to moderate the use of fossil fuel. Then, it would be necessary to trade the risks of nuclear accidents against those of climatic change. What then will be the general opinion? Undoubtedly, there will be many who protest that they do not care for the dilemma in which they find themselves, and that they would prefer that they and others should use energy from renewable sources, or that they and others should use less energy of any kind. But the general opinion is likely to settle for the lesser evil.

But the dilemma of the greenhouse is not as remote as may be thought. It is already something of a mystery that the atmospheric concentration of carbon dioxide has increased by a third without climatic fluctuations showing through the usual interannual noise. Most probably, the estimates of the changes derived from simple climatic models are overestimates. The fact that models have only recently been able to take reasonably realistic account of cloudiness, negative feedback in the context, is one possible explanation, but there are many other mechanisms that could account for a delayed response of climate to increased atmospheric carbon dioxide. While no one year's unusual weather can be statistically significant, it will be a great surprise if the greenhouse has not made its appearance before many decades have passed.

What will happen then? On the model of the Convention on the Protection of the Ozone Layer, signed last year, there will be international negotiations to limit the use of fossil fuel. There will be endless arguments about the magnitude of the limits and about the quotas due to various countries. But with luck, there will be a convention, and even the most reluctant communities will begin building nuclear power stations again. Shoreham itself may by then, of course, have been torn down, or perhaps converted to some other use. □

Making European policy

Last week's sensible NATO compromise points to the need for an external policy for Europe.

THERE are two predictably contrasting views of last week's meeting of the ministerial conference of the North Atlantic Treaty Organization (NATO). One is that the underlying disagreements were so great that important issues were buried in a frosting of sugary compromise, the other is that NATO's members triumphed over centrifugal tendencies by acknowledging that each must accommodate the views of others.

There is some support for both opinions, which are identical when allowance is made for the emotiveness of the language used. In deference to West Germany, as much concerned with what the opening to the East may hold as with the fear that modernized nuclear weapons, if ever used, are most likely to fall on German speakers, the pursuit of modernization was ringed



with caveats, which is wise. But the ambitions of the more hawkish members that there should be some counter-balance for the impending disappearance of intermediate missiles from Central Europe were neither embraced nor denied, but instead overshadowed by the more worthwhile resolution that there must be an attempt to negotiate in Europe an understanding about the balance of conventional forces in that densely populated region. That is not a compromise, but the only sensible course to follow at this stage (see *Nature* 332, 1; 1988).

But how? It is easier for a government to resolve that there should be an accommodation with a potential adversary than to ensure that negotiations will actually take place, and then succeed. Recent experience in this field is not encouraging. After nearly a decade, the East-West talks in Vienna on the balanced reduction of conventional forces are being abandoned, while the security talks under the aegis of the Helsinki agreements, while useful in reaching agreements such as those which now enable eastern and western observers to inspect each other's military manoeuvres, have little to say about the numbers of tanks and strike aircraft facing each other in Central Europe.

If another effort is to succeed, it will have to begin from painstaking and agreed inventories by East and West of their reciprocal conventional strength (with short-range nuclear weapons included). That is how the Euromissiles treaty (INF) was negotiated, and how the more ambitious treaty on strategic weapons may yet be ready in two months or so. The snag is that mere numbers, always the most easily estimated of quantities, are more relevant in estimating strategic than tactical strength, where geography and environment may be crucial.

Meanwhile, even those who take a cheerful view of last week's events should watch out for more distant but now predestined changes. Both in Europe and the United States, there is now a fashion for making gloomy parallels between the decline of the British Empire or even the Roman Empire with the supposed decline of the power of the United States. The argument is outrageously exaggerated, but there is nevertheless a sense in which European NATO members had better plan to play a more active part in their own defence.

Quite apart from imponderables about the future influence and (more important) inclinations of the United States, there has been a pronounced shift in the relative economic power of NATO's two halves in the past thirty years — in Europe's favour. US congressmen's complaints that Europe should do more to provide for its own defence are not mere sourness. It is also true that, with the INF treaty signed, the US contribution to the defence of Western Europe has been diminished in precisely the respect it is uniquely qualified to contribute. But the European members of NATO, now more anxious than ever to have an important say in the development of their relations with the East will find it simplest to pay for it.

This does not imply that Europe, having seen the back of exogenous long-range missiles, must arm itself to the teeth by other means. Nor does it follow from the pursuit of a strong defence of Western Europe that present and future improvements of political relations between East and West will be jeopardized.

Nobody suggests that Europe should transformed overnight from an armed camp to a demilitarized zone. Moreover, there is a need that Western Europe, with its present ambitions to become (from 1991, if the optimists are taken at their word) a properly unified community, must face the neglected problem of how a loose federation such as it will be should devise an external policy for itself. The European Commission at Brussels may have come a long way since the Treaty of Rome was signed, but it is still much better at drafting statements on external affairs than at making policy. The danger in the way in which the wind is at present blowing is that Western Europe, even if persuaded by events (and the United States) to spend more on conventional defences, will be poorly defended for lack of a wider framework in which to plan. □

Running in the dark

The present phase of the presidential election in the United States is disheartening.

IN the arcane catalogues of national customs, the procedures by which people elect their leaders deserve to be recognized as the most idiosyncratic. The Republic of Ireland, for example, is famous (among other things) for a system of proportional representation most outsiders cannot fathom, but which apparently ensures that no government will have a substantial majority over its opponents. The United States, by contrast, has devised a fiendish system whose chief effect seems to be that voters, before being given an opportunity to choose their next president from a list of candidates, are treated to the spectacle of members of the two principal parties abusing each other vigorously, but hardly bothering to complain at their titular opponents' policies.

That, on this occasion more than most, seems to be the consequence of the US system of primary elections, ostensibly a sensible way of enabling like-minded people to decide among themselves who will be their champion in the election that really matters. The result is that, since the beginning of the year, two dwindling bands of office-seekers have been shifting camp from one rural part of the United States to another, most of all concerned to undermine the reputation or the promise of their fellow Democrats or Republicans, as the case may be. What this does for the reputation of politicians is anybody's guess. And, in spite of the way in which a score of mostly southern states coordinated their primary elections on Tuesday this week in the hope of acquiring more influence on the outcome, the pollsters expect the abuse to continue until the summer.

While there may be few who fail to derive a little pleasure from the discomfiture of normally imperturbable people, the pity on this occasion is that the would-be candidates have allowed their obsession with their rivals' defects to dampen their participation in what might have seemed the purpose of the primary elections, that of giving them a chance to say how, if eventually elected, they would tackle important issues, of which there are many. Instead of policies, there are mostly only indications: Senator Robert Dole (Republican) and Representative Richard Gephardt (Democrat), for example, talk more openly about solving economic problems by protectionism, but not so specifically that voters would be able to complain if either became president and then followed quite different policies (as, it must be hoped, he would). And although the roster of prominent candidates in the primary elections includes one (Democratic Senator Albert Gore) with a distinguished record as a legislator concerned with science and technology, both in the House of Representatives and now the Senate, little has been heard of how research may help solve contemporary problems.

Part of the problem is the ghost of Mr Walter Mondale, the Democrat who stood for the presidency in 1984 and who lost. Legend has it that Mr Mondale's defeat is attributable to his ready declaration that, if elected, he would increase taxes so as to balance the US budget deficit (and also to finance various social programmes). In reality, there may be other explanations. But the result has been that none of this election's hopefuls still in the race has faced up to the budget problem, arguably the most serious for the new man's first year in office. Instead, there appears to be a conspiracy among the runners (in which the administration has joined) not to worry about the budget deficit expected this year (which may amount to \$175,000 million) and the accompanying trade deficit (\$160,000 million?). But that is merely the immediate difficulty. The present administration believed it would avoid the problems it has created because rising prosperity (and faster-rising taxes) would pay for the rising costs of an ageing population. That miscalculation points to a structural imbalance in the system that will make things worse before somebody has the courage to make them better. □

NIH gearing up to tackle the human genome

- New office to be created
- Prospects for 50 new fellows

Washington

THE National Institutes of Health (NIH) have left the sidelines and entered the fray in the effort to map and sequence the human genome. Director James Wyngaarden last week revealed NIH's plans following a two-day meeting of an *ad hoc* committee on complex genomes that he brought together to advise him on how NIH should proceed. The meeting marked a change in course for NIH and the human genome project. The recommendations did not consider whether the project should proceed, but how.

Wyngaarden says budget commitments from Congress and the Reagan administration have now made it possible for NIH to formulate long-range plans. Congress appropriated \$17.3 million to NIH specifically earmarked for genome activities in 1988, and the president's 1989 budget requests \$28 million for the project.

The first step will be to form a new genome office in the office of the NIH director, headed by an associate director. Although money will still be doled out by the individual institutes (primarily the National Institute of General Medical Sciences) the new associate director will co-ordinate the grant-making process, guided by a programme advisory committee.

NIH's plans fall largely along the lines recommended by the National Research Council (NRC) in a report released last month (see *Nature* 331, 467; 1988) with total spending set ultimately to reach approximately \$200 million per year. An immediate goal is a genetic linkage map with a resolution of one centimorgan. A working group of the *ad hoc* committee recommended that such a map, constructed with generally available probes, should be completed "as rapidly as possible". The working group estimated the cost for completing a linkage map in the next 3–5 years at \$50–70 million.

A key aspect of the NIH efforts will be graduate and postgraduate training. It is recommended that NIH should support 50 new post- and pre-doctoral fellows to ensure a steady supply of researchers needed to complete the project.

The basis for making grants will be projects that are intended to improve the scale of current activities between five- and tenfold. This idea of using improvements in scale as a basis for making grants is found frequently throughout the NRC report. It is intended to keep a rein on the project, making investigators provide

specific targets for improvements. Maynard Olson of Washington University and a member of the *ad hoc* committee says that loose definitions make it difficult to judge a project's success or failure.

Olson, who played an instrumental role in developing yeast artificial chromosomes that hold the promise of creating DNA clone libraries of 1 million base fragments, says the scale concept is crucial to make sure the project does not get out of hand. Olson says that although there is some fear that a large mapping and sequencing effort would rob other projects of funding, another concern is that the project would be propelled for political rather than scientific reasons.

As one of the few researchers involved in projects aimed at physical mapping of entire genomes, Olson worries that others have underestimated the difficulties of making such a map of the human genome. The largest genome to have a high-resolution physical map is *E. coli*, but that is only 4.7 million bases, compared to 3,000 million for the human genome.

Completion of the project does not consist solely of establishing a sequence for the human genome. Wyngaarden has made it clear that analysis of the genomes of other species will be an integral part of the project. Wyngaarden also says that NIH will continue to support individual investigations designed to make sense of sequence information as it is amassed.

Reaction to NIH's plans has been mostly enthusiastic. James Watson, director of the Cold Spring Harbor Laboratory, says he is extremely pleased that NIH have taken the bit in pursuing the genome project. David Smith, who heads the Department of Energy's efforts in this area says now that NIH have staked out a position it will be easier to establish momentum for the project.

George Cahill, of the Howard Hughes Medical Institute, also a participant in the *ad hoc* meeting, says that Hughes' efforts will focus on data handling. Work on merging genetic databases will begin immediately while NIH are putting their plans into practice. Congressional interest in the genome project will also be a factor in how far NIH can go with it. Congress may be reluctant to put NIH in charge of what will certainly be a technology development programme in its early stages. That role may belong with the Energy Department and the National Science Foundation.

Joseph Palca

Congress keen to top NIH's requests

Washington

"By far the best budget proposal from any administration in a decade" is how NIH director James Wyngaarden described the 5.4 per cent budget increase proposed by the administration for 1989. But Wyngaarden, testifying before the House Appropriations Subcommittee in his role as spokesman for the administration, did not deter his traditional supporters. The subcommittee appears ready once again to give more than NIH is officially asking for.

Wyngaarden's testimony focused on the achievements of the NIH and contained no hint of a hard-luck story. But he did voice worries, shared by members of the committee, that NIH are losing good people to universities and teaching hospitals because of large pay differences. "It's not yet a haemorrhage but remedial action is needed", he claimed.

In cross-questioning by the committee, the adequacy of the budget came under close scrutiny. For in spite of the "magnificent increase" heralded by Wyngaarden, NIH's budget formulation office estimates that a further \$500 million is necessary to maintain NIH's current levels for 1989.

For grant applicants, the biggest worry is that there will be 800 fewer grants next year. The level of funding per grant is also at issue. Although rising since 1982, the level of 1974 has only recently been regained. And the discrepancy between the awards recommended by the study section reviewers and the amounts actually awarded (downward negotiation) will reach a record 13.3 per cent in 1989. Meeting the targets for numbers of grants means less to spend on each.

Committee members pressed Wyngaarden to identify projects and opportunities that would be short-changed or excluded under the proposed budget. This yielded few specific answers apart from the clinical trials programme where funds for new trials are severely limited. Wyngaarden promised to submit details of what had been cut from a much larger, provisional request prepared in an earlier phase of the budget negotiations.

Subcommittee chairman William Natcher (Democrat, Kentucky) seemed confident that Congress could outdo the administration.

This will be a "People's Appropriation Bill" he claimed, arguing that money for such a good cause would not be refused. Some of his staff are less sanguine. The Bipartisan Budget Agreement, which limits the overall increase on non-defence spending to 4.8 per cent, has placed a ceiling on everyone's expectations.

Penelope Austin

US Congress hears the good news on Chinese food

Washington

A US congressional hearing is not usually the place to hear the latest scientific results. But last week, those attending the Senate Committee on Governmental Affairs had a preview of a massive study on the relationship between diet and disease carried out in China by Cornell University, the University of Oxford and the Chinese Academies of Preventive Medicine and Medical Sciences.

From the highlights presented by T. Colin Campbell, professor of nutritional biochemistry at Cornell University, the Chinese study seems likely to provide a massive database for testing the relationships between diet and disease and to affect significantly the nutritional advice given to Westerners.

The Senate committee was taking evidence on "Nutrition in Health and Disease" with the aim of clarifying "the confusion spawned by the proliferation of dietary advice" and finding a more effective federal government role in nutrition research. The first-ever Surgeon-General's report on nutrition and disease will be released this summer.

Campbell's view is that nutritional advice has been too fragmented. Recommendations have focused on single nutrients without a wider perspective that takes into account the enormous complexity of biological relations.

The only way to sort out the complex interactions of diet, disease and human variation is through examination of large cohorts of people, with stable but diverse eating habits. China is perhaps the only country in the world where this can be easily done. Rural Chinese people usually remain in one area all their life, eating the same kinds of local foods. Incidence of disease and dietary habits vary enormously. In contrast, the mobile populations of the west have become more uniform.

China also has the organizational capacity to carry out large-scale surveys and an epidemiological tradition to live up to. In the 1970s, a massive epidemiological survey was carried out that resulted in the now famous *Atlas of Cancer Mortality in the People's Republic of China* (Nature 318, 218; 1985). The survey charted striking differences in the incidence of various cancers in different regions of the country.

That atlas forms the base for the new survey linking diet and disease. Sixty-four counties, selected for their extreme variation in diet and in mortality characteristics, were revisited by Chinese teams. Over a five-year period, blood and samples and information on lifestyles were collected from 6,500 people in 130 villages. Half of the 6,500 also provided urine samples and

day-by-day details of their diet. Food samples were taken away for analysis. Altogether, the teams put in some 600 man-years of professional labour.

Specimens then went out to laboratories around the world. Data were assembled on hormone, mineral and fat levels; on the incidence of virus antibodies; on dietary composition; on individual health and many other factors. A total of some 350 variables were recorded.

The results are now sitting in a computer at the laboratory of epidemiologist Richard Peto at the University of Oxford. Through the generosity of the Chinese, whose efforts Campbell estimates would have cost "more than \$100 million" in the United States, the world's researchers will be given free access to the databank later this year.

A few nuggets have already been mined from the data. One surprise is that Chinese people consume about 20 per cent more calories than Americans. But there is very little obesity among the Chinese, showing that calories alone do not cause excess weight and are unlikely to determine chronic disease. The Chinese do not eat much fat: on average, fats make up only

15 per cent of calorific input, less than half the amount usually recommended in the United States. But the low fat intake does not seem to produce any ill-effects.

Chinese people have serum cholesterol levels that are so low that they are almost off the bottom of the scale for the US population. Cardiovascular disease is correspondingly low, as expected, but worries raised by US data showing an increased risk of colon cancer at very low serum cholesterol levels appear groundless.

The Chinese diet is also full of fibre: average intake is 34 grams a day, more than three times the US average but close to the amount recommended by the National Cancer Institute. Fears that high-fibre diets could hamper mineral uptake are not substantiated. Plasma, iron, zinc and magnesium are all at healthy levels.

Although the Chinese study is the biggest collection of diet and disease data yet put together, it is still small compared with what Campbell has planned. Additional surveys will soon add more detailed mortality statistics for the areas where the diet-disease survey was conducted. In 1989, a supplementary survey will be launched. And beyond that, a prospective study of 525,000 individuals is being planned. That should provide data to keep researchers busy for the next 50 years.

Alun Anderson

Brighter travel prospects for Hungary's hard-up scientists

London

RECENT cuts in the Hungarian science budget and the new relaxed passport laws could lead to more Hungarian scientists taking jobs abroad. Since 1 January, every Hungarian citizen has an automatic right to a passport, although to travel to the West, an appropriate amount of hard currency is also necessary. Although there seems no immediate likelihood of a brain drain, there will almost certainly be a significant increase in dependence on the facilities of foreign laboratories to do experimental work, a feature of Hungarian scientific life in recent years.

Scientific exchanges and visits are recognized as a stimulus to scientific success. But many of the Hungarian scientists who make the greatest use of Western facilities are embarrassed that such visits give access to apparatus Hungary cannot afford to buy.

Last July's reform budget considerably exacerbated the situation. The central budgetary support for science was cut by some 15 per cent over 1987-88, but taking inflation into account, the real drop in funding will be about 30 per cent.

Dr Ivan T. Berend, president of the Hungarian Academy of Sciences, has expressed a fear that some academy insti-

tutes and academy-supported university research would have to close has so far not materialized, partly due to the former system of centralized economic management. Under this now discredited system, there was a sharp line of demarcation between research and development institutes and manufacturing enterprises. Although for some years in-house research has been permitted and even encouraged, Hungarian industry has, by no means, taken over responsibility for the research and development base, and, as the available central funding has fallen off, research institutes have secured more and more direct research and development contracts from industry and have even begun running small production projects on their own account.

But the resources and demands of Hungarian industry are finite, and it is particularly difficult for basic research to secure appropriate contracts from the industrial sector. Fortunately, the 4,000 million forints one-off emergency injection of support for basic research which the government authorized in 1986 has not been affected by the cuts and will, in the short term, help to cushion the effects of the recession. But the long-term prospects are less bright.

Vera Rich

Future pattern of UK earth sciences starts to emerge

London

EARTH science departments in British universities will be classified into three types, with differing provision for teaching and research, if long-awaited recommendations announced last week are adopted. A comprehensive review of earth science departments, commissioned by the University Grants Committee (UGC), proposes the following classification of departments.

Type M. Medium to large departments providing teaching and research in mainstream earth sciences, including single honours courses. Some of these departments would run major items of equipment that would be made available to the wider community.

Type I. Medium to large departments or departmental groupings in which earth

'Group 1', which will be given greater student numbers than other department of the same type (Group 2). It is recommended that the distinction between Groups 1 and 2 should not be permanent: the reorganization would be monitored at regular intervals by external visiting groups, which would inform the institutions and funding bodies of a department's progress, and would be empowered to recommend changes in status.

In the proposed scheme, all but two of the existing earth science departments have been classified (see table), with Reading and, more curiously, Bristol having yet to learn their fate.

The main surprises in the recommendations are that Imperial College has been classified as Type I rather than M, Swansea is to lose its department entirely, and Sheffield and Hull are to retain a provision for earth sciences only through other departments.

The national review committee was also asked to consider the Open University on the same basis as the institutions within the UGC's purview and decided that the university's earth science provision was equivalent to Type M, Group 1.

While earth scientists are generally relieved that the model proposed by Professor Ron Oxburgh for a more closely defined three-tier system (see *Nature* 326, 813; 1987) has been considerably watered down, and that the present model appears to incorporate a welcome degree of flexibility, there is concern that the level of concentration proposed in the new system will result in a net loss of opportunity for students to read earth sciences. Despite more widespread relief that the principal recommendations have been made, the review remains far from complete. Questions still need to be answered about the resources available for the whole reorganization and the allocation of resources between institutions. The joint UGC/NERC (Natural Environment Research Council) equipment rationalization committee, which will recommend where major pieces of apparatus should be housed, has yet to report.

A further uncertainty remains over the employment status of staff transferring between institutions. The government is presently trying to pass legislation that would abolish academic tenure for university staff taking up new appointments since last November, when the legislation was proposed. The Committee of Heads of University Geoscience Departments wrote to the government in January to try and obtain exemption for staff forced to move under the UGC reorganization, but still awaits a reply. **Simon Hadlington**

MoD's sums wrong?

THE British government has for years been overestimating the nation's total investment in research and development, according to a study by the Society for British Aerospace Industries (SBAC).

The study, which involved several of SBAC's larger member companies, concluded that Britain invests no more than 1.9 per cent of its gross domestic product in research and development, compared with the 2.3 per cent quoted in government statistics. The study was undertaken when several defence manufacturers found it impossible to reconcile the value of work in their Ministry of Defence (MoD) contracts with the ministry's published expenditure on research and development. SBAC says that MoD accounting conventions, rather than being concerned with identifying funds it spends on research and development, are aimed at controlling the costs of projects. "But this can result in work which has little or no innovative content being categorized as R and D when it should not be counted as such", says SBAC. **S.H.**

Chinese prosper

ALL Chinese scientists now have the right to get rich in their spare time. Following a successful pilot scheme, the Chinese government has decided to allow scientists and technical personnel to take second jobs for extra income. Scientists can now offer consultancy and advisory services to local farms and enterprises but they may have to share their fee with their main employer if his rights are affected. **V.R.**

Antarctica-on-Sea

ALREADY forced to deal with difficult logistics and tight budgets, the US National Science Foundation (NSF) Antarctic programme has a new problem on its hands: tourists. NSF estimates that 7,200 people will visit the Antarctic this year, double the number of government research and support staff supported by the 18 countries conducting research at the pole. NSF is talking to tour operators, hoping to minimize the impact on research, but it is accepted that the numbers of visitors is bound to grow. **J.P.**

Superhoax at 190 K

A REPORT in the London *Times* last Saturday that scientists at the Cambridge University Research Centre in Superconductivity had a 190-K superconductor turns out to be the result of an undergraduate hoax. The success of the prank, which was perpetrated as part of the university's "rag week" celebrations, underlines the continuing popular interest in the race to higher superconducting transition temperatures. **L.G.**

Earth sciences departments

Type M Type I Type J

Group 1	Cambridge	East Anglia	
	Edinburgh	Lancaster	
	Leeds		
	Liverpool		
	Manchester		
	Oxford		
Group 2	Birmingham	Aberystwyth	Aberdeen
	Cardiff	Imperial	Keele
	Durham	Southampton	St Andrews
	Glasgow		
	Leicester		
	RHBC*		
	UCL/Birkbeck		

* RHBC, Royal Holloway and Bedford New College.

Earth sciences provision within other departments

Exeter	Nottingham
Hull	Sheffield
Newcastle†	

† Special collaborative arrangements will apply between Newcastle and Durham.

No on-site provision

Aston	Strathclyde
Dundee	Swansea

scientists would contribute to interdisciplinary activity as well as mainstream earth sciences.

Type J. Small to medium departments or groups offering joint honours and service teaching, but without expensive earth science equipment on site.

Additionally, there would be institutions with no earth science departments as such, but which would provide teaching in some aspects of earth sciences located within an appropriate departments.

Within Types M and I, departments seen as having particular strength and potential have been further categorized as

Japan's frontiers research to cross international boundaries

Tokyo

JAPAN'S Human Frontier Science Program is at last taking shape but it is considerably smaller than originally planned. Nevertheless, if Japan can win support from the governments of the other Western summit members (Canada, Britain, France, West Germany, Italy, the United States and the European Community), Frontiers could become a substantial international programme of research in the biosciences.

In 1984, the then Japanese Prime Minister, Yasuhiro Nakasone, first suggested that Japan should initiate an international bioscience project. Two years later, the Ministry of International Trade and Industry (MITI) and the Science and Technology Agency (STA) came forward with proposals — the Human Frontier Program and the Human and Earth Science Program — which were later fused into the Human Frontier Science Program.

It was widely reported that Frontiers would plough about \$6,000 million into international basic research in the biosciences over twenty years. But the programme has been bogged down in a prolonged feasibility study and the scale of the programme has been reduced.

Last week, scientists and science administrators from the seven summit nations and the European Community met in Tokyo to finalize conclusions of an "international feasibility study" which began in November. The committee in a draft report recommends that Frontiers should consist of international grants, fellowships and workshops. A council of about twenty scientists will select priority areas for research and topics for workshops, while review committees will screen applications for grants and fellowships.

Initially there will be two priority areas for research: "higher order functions of the brain" and "molecular recognition and response functions". According to the committee's recommendations, 30–50 three year grants each worth an average of half a million dollars a year will be awarded annually to teams of researchers from two or more member nations. In addition, 100–200 two-year postdoctoral fellowships worth about \$50,000 each a year will be awarded annually and 10–20 workshops will be held every year at a cost of about \$100,000 per workshop.

This brings the total annual budget for Frontiers (excluding administrative costs) to \$60–100 million. But it is still not clear where the money will come from.

MITI officials say they now have the "understanding" of the Ministry of Finance and a programme of the scale proposed is possible. But the other summit nations have yet to make any financial

commitment to Frontiers and it would be unprecedented for Japan alone to finance an international programme of this scale. MITI and STA have only won a few million dollars for the programme in fiscal year 1988. But MITI is in a strong position to bargain for funds in the next fiscal year. The Finance Ministry, Prime Minister Noboru Takeshita and the ruling Liberal Democratic Party are keen to introduce a new indirect tax and MITI's support will be required to gain the "understanding" of Japanese businessmen. But no doubt some commitment on the part of other summit nations to join Frontiers would greatly strengthen MITI's hand.

Another critical test facing the pro-

gramme is the selection of scientists for the governing council and review committees. Sir John Kendrew, representing the European Community, proposed in last week's meeting that the seven summit nations and the Community should each nominate five scientists for the council. And then Japan as initiator (and financier?) of Frontiers should choose twenty of them. Although well received, this proposal would give the Europeans a disproportionately large say in the programme.

Time is running short. MITI and STA are keen that an international peer review system should be in place to distribute grants and organize workshops in fiscal year 1988 (which starts next month). The next Frontiers meeting will be held in Bonn on April 29 and the ministry and agency are clearly aiming at getting the programme off the ground before the next summit in Toronto. **David Swinbanks**

Low-energy accumulator for the French Saturne synchrotron

Paris

SATURNE, a French synchrotron with a unique capability to accelerate polarized protons and deuterons, is now equipped with a low-energy accumulator booster synchrotron, MIMAS. Operational since October 1987, MIMAS was officially unveiled on 2 March at the Nuclear Research Centre (CENS) at Saclay, south-west of Paris.

Scientists at Saclay hope that MIMAS will allow Saturne to accelerate ions as heavy as krypton, as well as increasing the intensity of the polarized beams. After accumulation and acceleration, MIMAS can inject a particle beam at up to 47 MeV for protons (11 MeV per nucleon for heavy ions) into the main synchrotron,

Saturne, where energies up to 3 GeV for protons and 1.1 GeV per nucleon for heavy ions may be achieved.

"The problems of working at such low energies have scared off most laboratories", says Michel Oliver, of the Laboratoire Nationale Saturne (LNS) at Saclay. The principal problem, beam contamination by unwanted electrons, is overcome in MIMAS by heating the vacuum chamber to 300 °C and trapping electrons on a titanium sublimator.

Although the synchrotron is unable to attain the high energies necessary to work with heavy ions such as uranium, the wide range of available particles and the high quality of Saturne's spectrometers means that it is much in demand. Last year, out of 400 scientists using the facility, 270 came from abroad, with roughly the same number of projects having to be turned down. A cooperative research agreement has recently been signed with the GSI Laboratory at Darmstadt, West Germany, where a heavy-ion synchrotron is being built. A similar exchange is under discussion with the Italian National Nuclear Physics Institute.

The £3.7 million budget for Saturne is financed jointly by the French Atomic Energy Commission (CEA) and the Centre National de la Recherche Scientifique (CNRS). Foreign researchers do not have to pay for beam time in the laboratory.

MIMAS has provided another kind of spin-off. In collaboration with an industrial company, GEPE, the laboratory at Saclay has recently patented the high-precision electronic relay it invented as part of the technological development for the new synchrotron. **Peter Coles**



Isolation supports of Saturne's polarized particle source.

Superconducting conference yields new temperature record

Interlaken, Switzerland

THE record for high-temperature superconductivity has been broken yet again. On Wednesday night last week, Dr Paul Grant of the IBM research laboratory in Almaden, California announced the attainment of zero resistance at 125 K in a thallium-barium-calcium-copper oxide. At the time of the announcement, this represented an increase of over 20 degrees on the highest published zero-resistance temperature.

Grant was speaking at an evening session of the week-long International Conference on High-Temperature Superconductivity and Materials and Mechanisms of Superconductivity, which brought 1,200 people from 36 countries to Interlaken, mostly to hear reports of the latest research on the high-temperature oxide superconductors. Of the 800 talks and posters, more than 750 reported work on the oxide materials.

Most of the excitement was generated by reports on the new rare-earth-free bismuth and thallium superconductors. Superconductivity at over 100 K in bismuth-strontium-calcium-copper oxide had been reported by Hiroshi Maeda less than six weeks before the conference began (see *Nature* 331, 377; 1988), yet more than 30 speakers were able to report results on this system. Only a few had had time to work on the thallium system, discovered by Z.Z. Sheng and A.M. Hermann of the University of Arkansas only in mid-

February (their report of the discovery is published on p.138 of this issue), but by the end of the week Z. Zhao of the Institute of Physics in Beijing was able to announce that his group had measured zero resistance at 120 K, almost equalling Grant's record. The frenetic pace at which work is proceeding around the world was evident as speakers presented data that were only hours old, hot off the facsimile machine.

Although the race to ever-higher transition temperatures lent spice to the atmosphere, it did not overshadow the equally important scientific advances reported on the new materials. It was generally agreed that there are two closely related superconductors in the bismuth family, with stoichiometry near $\text{Bi}_2(\text{Sr}, \text{Ca})_2\text{Cu}_2\text{O}_8$ and nominal transition temperatures of about 80 and 105 K. Several groups have determined the crystal structure of the lower-temperature phase and find that it resembles the 40-K lanthanum and 90-K yttrium superconductors in containing perovskite-like layers with extended planes of copper and oxygen. But the linear copper-oxygen 'chains' thought by some to play a role in the elevated transition temperature of the yttrium compound are absent in the bismuth structure, which instead has double layers of bismuth oxide.

The absence of chains came as an optimistic sign to many who, a year after Paul Chu's discovery of the yttrium superconductor, had begun to wonder whether there were any more high-temperature superconductors to be discovered. As R.J. Cava, of AT&T Bell Laboratories, said at the conference, "The yttrium structure, with its copper-oxide chains, was very special — almost a one-off. But just about every compound with copper and oxygen has planes." The discovery of the thallium superconductors, which appear to be structurally similar to the bismuth compounds, reinforces the view that the field is now wide open for further discoveries. In fact, as thallium is highly toxic and bismuth is in short supply, both materials may be of greater scientific than technological importance.

Although the new materials stole the limelight, the conference was also notable for the large volume of high-quality work reported on the now suddenly 'old' lanthanum and yttrium materials. Steady progress was evident on the materials side, as critical currents in both thin-film and bulk material continue to increase. But attempts to understand the mechanism of the superconductivity are still frustrated by the prevalence of conflicting experimental results, largely owing to

Indian science budget

New Delhi

INDIA'S annual budget for 1988–89 provides Rs17,000 million (£850 million) for scientific research and development. This is 6 per cent of the total and 20 per cent more than the scientific agencies spent last year. In comparison, defence gets a whopping £6,500 million, of which £320 million is for military research (see table).

As in the past, the Departments of Atomic Energy and Space between them share 70 per cent of the budget. The Department of Atomic Energy is setting up a major research centre for accelerators and lasers at Indore in Madhya Pradesh, and the Space Department plans to build a multi-million dollar facility for developing cryogenic engines for its planned geostationary launch vehicle.

Spending on biotechnology and microelectronics is to increase and £4 million has been ear-marked for work on high-temperature superconductors. Prime Minister Rajiv Gandhi is himself chairman of the Apex committee that is directing superconductor research. New facilities for which provisions have been made include a medium-range weather forecasting network using the supercomputer being imported from the United States, a pilot plant for manufacturing amorphous silicon for solar cells and a second permanent station in Antarctica.

Gandhi's budget in a drought year is predictably biased in favour of farmers and the rural poor. By exempting taxes on electronic items made in villages, he hopes to push industries to rural areas and create jobs.

K. S. Jayaraman

	Rs (millions)
Atomic energy	8,490
Space	3,540
Agriculture research	960
Electronics	1,400
Ocean development	220
Biotechnology	450
Industrial research	880
Medical research	200
Other	930
Defence research	6,400

variable sample quality.

Two years ago, the first conference on Materials and Mechanisms of Superconductivity in Ames, Iowa was attended by about 250 people; the record superconducting transition temperature then stood at 23 K. At Interlaken the corresponding figures were over 1,200 people and 125 K, leading T.H. Geballe of Stanford University to suggest that conference attendance grew in proportion to transition temperature. As Geballe has been charged with responsibility for the next conference in the series, to take place in the summer of 1989 in the San Francisco area, he has mixed feelings about the prospect of room-temperature superconductivity.

Laura Garwin

Thin film already

Tokyo

ALTHOUGH bismuth-strontium-calcium-copper oxide (see *Nature*, 331, 377; 1988) superconductors were discovered barely a month ago, a Japanese company has already succeeded in making a thin film of the new material. The same development took many months for the rare-earth superconductors.

Matsushita Electric Industrial Co. announced last week that it has made a polycrystalline thin film that superconducts at 100 K, at least ten degrees higher than reported for films of yttrium-based copper oxides. Although the increase in critical temperature is modest, it does double the operating margin above 77 K, the temperature of liquid nitrogen.

The film has a critical current density of 10^5 A cm^{-2} , about half that reported for polycrystalline rare-earth thin films, but Matsushita researchers are confident that they can increase current-carrying capacity by preparing monocrystalline thin films.

David Swinbanks

South African universities are squeezed by still more cuts

London

SOUTH African universities seem more than ever at loggerheads with the government after the events of the past month, in particular the announcement that government payments to universities will be reduced by between a further 8 and 12 per cent in the academic year now beginning.

This decision followed by only two weeks the publication of a report by the Committee of University Principals (CUP) pleading, among other things, that the funds generated by the subsidy formula should be passed on in full to the universities. The proposed cuts are distinct from the threatened discriminatory cuts of the subsidies to universities which have displeased the government.

The committee responsible for the CUP report, under Professor J.P. de Lange (who resigned in 1986 as rector of the Rand Afrikaans University), is likely also to be disappointed in its plea that university salaries should be urgently increased. Its report also asks for a rationalization of the university system to cope with dwindling resources, the creation of a single ministry for tertiary education and the study of new ways of identifying students likely to benefit from tertiary education.

All universities will be hurt by the reduction of subsidies, which are determined by a complicated formula for calculating an institution's 'ideal' need, based on graduate and undergraduate student numbers (differently weighted), success rates and research output. Universities are expected to pay a proportion of this amount from fees paid by students and from endowment income. The balance of the universities' budget is paid by the state, which has cut its contribution by 17 per cent per annum over the past 3 years. Now the government has announced that in 1988 this reduction will be increased to between 25 and 29 per cent.

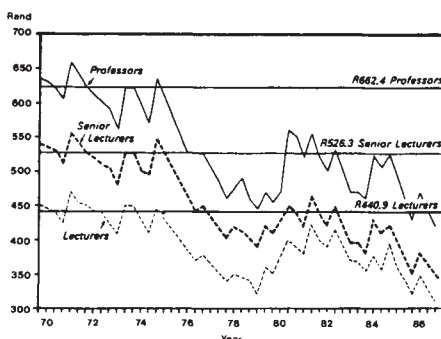
Apparently reconciled to continuing privation, the report recommends that CUP should put forward guidelines for rationalization of the 21 universities. It suggests a system of accreditation to determine which departments have the right to accept graduate students rather than one in which whole institutions will be restricted to undergraduate teaching.

Academic salaries are another bone of contention. The CUP report says that, in real terms, professors' salaries have declined by almost 50 per cent since 1971, and are now the same as those of lecturers in 1981. And while inflation is now running at 15 per cent a year, academic salaries have been neglected except in medical faculties.

Despite the difficulties, the South

African university system appears to keep growing. There was an almost threefold increase of staff numbers between 1965 and 1985 while, during the same period, student numbers increased from 59,000 to 215,000. Overall, university students amounted to 7.7 per 1,000 total population in 1985, compared with 7.5 per 1,000 in Britain. For the white population, the proportion was 29.2 per 1,000, on a par with France, with the highest *per capita* student population in Europe. The CUP report notes the increasing participation of women, but a sharp decrease in the proportions of students reading engineering, science and medicine.

The CUP also records a rapid growth of non-white student numbers in all tertiary education institutions. Whereas white students accounted for 67 per cent of the total in 1979, the proportion will have fallen to 35 per cent by the end of this decade, and may be only 12 per cent by the



Real earnings (1970 prices) of professors and lecturers.

end of the century. Because of the decline of white births in the early 1970s, the admission of black students to the ten universities under the "white" Department of Education and Culture will increase, especially at the four universities with open admissions policies, so that the supervision of tertiary education by five ministries is more ludicrous than ever. The government has sought to control the growth of black numbers at the English-language universities by pegging the numbers qualifying for the state subsidy at 1986 levels.

On admission requirements, the CUP committee says that a candidate's potential as well as scholastic performance should be a consideration in the award of a university place. It also asks for an improvement of standards of secondary schooling and, cautiously, recommends a unified examining body at the secondary level. Recognizing that this proposal is unlikely to be acceptable to the government, the committee advocates bridging courses to counter inequities in primary and secondary education. **Michael Cherry**

French medics to go on-line

Paris

MEDLINE, the world's largest bibliographical database, is now integrated with the French national videotext network, Minitel. This will mean that MEDLINE's 4.7 million references can be consulted through any of the 3 million conventional Minitel terminals in France.

Potential users of MEDLINE will have to subscribe, but 'kiosk' access — the highly successful 'pay-as-you-go' charge system developed by the French telecommunications service for Minitel users — will soon be available. As users will also be able to order a photocopy of a desired article or its abstract without going to a terminal at a major library, it is hoped that the new service will encourage medical practitioners to keep up with current research.

The new MEDLINE service was launched in Paris last month at a two-day conference organized by INSERM, the national institute of health and medical research. The conference concluded a series of 12 national 'mini-symposia' exploring existing and possible future links between basic and clinical biomedical research.

Since 1985, INSERM has organized annual national research seminars around a specific topic. The theme of clinical research was felt to be appropriate this year, not least because the application of fundamental research in molecular biology and immunology has been so obviously important in recent research on AIDS (acquired immune deficiency syndrome) and other infectious diseases. But French scientists, like many of their western colleagues, also have to justify their demands on the state purse.

Co-financed by the Ministry for Research and Higher Education and the Health Ministry, INSERM spends about 25 per cent of its £100 million budget on clinical research. But almost half of its 249 research units are in hospitals and 32 per cent of full-time researchers are medical doctors. INSERM research units have benefitted from several national research campaigns and so have been less affected by government reticence to fund basic research.

INSERM nevertheless feels it has to be seen to carry out applied research. When she took over the Health Ministry in 1986, Michèle Barzach asked INSERM for specific research initiatives on the major fatal diseases in France. "There is no question of undermining basic research . . . it is indispensable", said Barzach at the conference in Paris. But she is "counting on INSERM" to come up with research that has some direct clinical applications. **Peter Coles**

San Andreas drilling halted

Berkeley

A PROJECT to drill down to the roots of the San Andreas fault has become one of the first victims of the failure of the National Science Foundation (NSF) to win its expected budget increase.

The Cajon Pass drill-hole, the deepest scientific drill-hole in the United States, was intended to go down to 16,000 feet. But research will stop after the completion of the second of three planned stages, at a depth of 12,000 feet. The project was designed to test a controversial hypothesis that the San Andreas fault is under much less stress than is currently thought.

NSF funds for Deep Observation and Sampling of the Earth's Continental Crust (DOSECC), the non-profit consortium of 40 universities that administers the Cajon Pass project, were cut from an expected \$6 million to \$3 million after last autumn's White House/Congress budget summit forced a slowdown in NSF's expansion. The amount available is insufficient to complete this year's drilling or take measurements in the new segment of the hole, says Stanford geophysicist Mark Zoback, director of the Cajon Pass project. But NSF has diverted an additional \$1.8 million from other projects in its Continental Lithosphere Program, to allow completion of the second phase of drilling and measurements. This reallocation of funds means that the project can expect nothing next year and will remain inactive unless it is refunded in 1990, said DOSECC president, G. Arthur Barber.

Proof that the San Andreas is a low-stress fault depends on sensitive heat-flow measurements which measure friction in the fault. Such measurements become more reliable at greater depths, where they are less likely to be influenced by shallow faults in the area. If the low-stress hypothesis was proved correct, it would mean there must be friction elsewhere resisting the forward motion of the plates, says Zoback, perhaps at horizontal faults deep under the surface.

The importance of the project goes beyond the stress question, according to Barber. Drilling projects are essentially extensions of NSF's individual grant programme because they enable individual researchers to carry out a wide variety of experiments in the hole and on its core. About 40 researchers from 10 universities are involved in the Cajon Pass project.

Barber has now begun to seek corporate funding for DOSECC projects. Petroleum and mineral exploration companies are eager to get involved in the project, he says, because of the technical lessons it is providing in drilling and core handling.

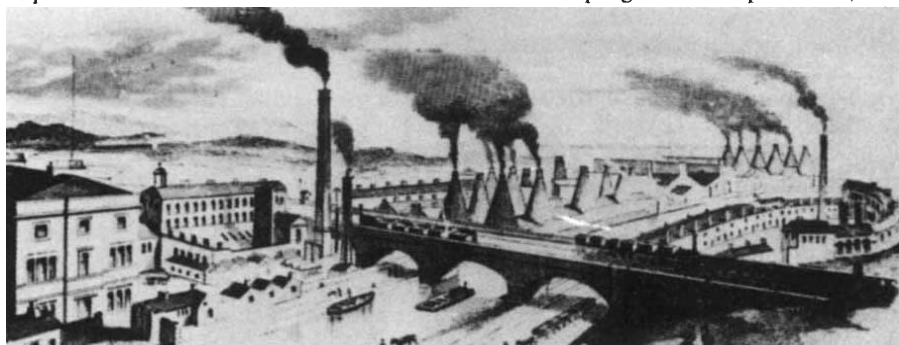
Marcia Barinaga

Transputer support centres will look for problems to solve

London

THIS week sees the official opening of the National Transputer Support Centres in Sheffield, part of a government-supported effort to turn microchip manufacturer Inmos into a money-maker instead of an example of British high technology wasted by an inability to apply and profit from an important invention.

Inmos at present is part of the entertainment and electronics group Thorn-EMI, which took the company over in 1985 and has invested heavily in transputer development. But Thorn-EMI is a consumer electronics group and is thought to be anxious to offload Inmos if a willing buyer can be found, despite the fact that Inmos is now at a stage where profits can be expected.



The Sheaf Works in 1855, when Sheffield's steel and cutlery trades were pre-eminent. The remaining building on this site, now a science park, is the first to be occupied by the national centre.

Applications for the transputer are the key to its future, hence the "Transputer Initiative" announced last autumn by the Science and Engineering Research Council (SERC) and the Department of Trade and Industry (DTI). This scheme will spend some £2.6 million over the next four years on setting up six Regional Support Centres with the brief to provide training and consultancy for industry, to promote research into transputer applications and to develop a software base for the exploitation of the transputer. For London and the South-East the centre will be at

SERC's own Rutherford Appleton Laboratory; for Northern Ireland both Queens and Ulster Universities are involved; the North West centre is at Liverpool; Scotland at Strathclyde; and the South West at Southampton.

The sixth centre, covering the North East and also taking on a national role in operating a software exchange library, was opened on Monday this week by the Minister for Higher Education and Science, Mr Robert Jackson. The national centre is a joint project of the University of Sheffield and Sheffield City Polytechnic, and the emphasis there will be on engineering.

What will these centres be trying to do? The uniqueness of the transputer lies in its combined programmable processor, ran-

dom access memory and communications channels on the one compact chip. The ability of the chips to communicate and process data at the same time makes transputers ideal for potentially powerful 'parallel processing' systems; in addition the transputer makes use of 'Occam', a computer language specifically developed for in-parallel operations.

So the transputer has another four years and £2.6 million to help it find a market, though Thorn-EMI may consider that too long to wait for a return on its investment.

Charles Wenz

Seven new 'specials' for Australia

Sydney

Australia is to open seven new "special research centres", bringing the total to fifteen. Called "centres of excellence" when the scheme began in 1982, the new institutions are intended to focus new institutions on research with potential economic benefit. A\$3.6 million will be spent on the centres this year. Among the recipients are Jack Pettigrew, who will head Queensland's first centre, for vision, touch and hearing. Pettigrew has been successful in exploiting basic research on animals to

develop a new binaural hearing aid.

Other centres are for human communication (Graeme Clark, Melbourne), protein and enzyme technology (Richard Wetterall and Robert Scopes, La Trobe), electronic structure of materials (Ian McCarthy and Erich Weigold, Flinders), lasers (James Piper, Macquarie), membrane and separation technology (Christopher Fell and Hans Coster, New South Wales), and industrial control science (Graham Goodwin, Newcastle).

Charles Morgan

Particle physics programme

SIR—Steven Weinberg's Commentary article (*Nature* 330, 433; 1987) presented the prevalent view that particle physics (defined as quantum field theory, general relativity and related areas of astrophysics and cosmology) is "at a level which is in fact in absolute terms quite deep, perhaps close to the final source". This optimistic view needs to be balanced by a more critical look at current problems in particle physics.

The 'Standard Model' of particle physics has not been able to retrodict, let alone predict, the masses of quarks and leptons or to provide an explanation for the organization of particles into regular families. The mass and interaction characteristics of the hypothetical Higgs boson, a key particle, cannot be predicted. Many variants on the particle physics theme require more dimensions than the familiar four of space-time, but there is no known way to test for them. The hypothesized unification of the four forces is predicted to occur at energy levels that will probably never be accessible to experiment.

Quark confinement, charm, the 'quark-antiquark sea', gluons, the Higgs mechanism and renormalization all 'work', but they are *ad hoc* theoretical fixes to contradictions between theory and observation. The hypothetical magnetic monopole is another key particle in the elementary particle pantheon, yet after 40 years of assiduous searching (and readjusting of parameters) physicists are still waiting for a real detection of a magnetic monopole. One wonders if particle physicists would allow nature to answer in the negative on this crucial issue. The fundamental prediction of the Standard Model, that spin would be irrelevant to high-energy elastic scattering, has been falsified by Kirsch's scattering experiments over the past decade.

In summary, there are at least 20 basic parameters that are crucial to the particle physics programme (such as particle masses, coupling strengths, magnitudes of CP-violations and so on) that cannot be rigorously derived and are therefore manually adjusted to agree with observations. Superstring theory, which has most particle physicists in a highly excited state, is completely untestable, for now and for the foreseeable future.

On the cosmological front, things are somewhat worse. In the standard cosmological model, the Big Bang (with or without inflation), the key events in the evolution of the Universe are supposed to have taken place within 10^{-25} s of the Big Bang, and thus are empirically unverifiable. None of the popular and much-touted models predicted the existence of the dark matter, the large-scale streaming of galaxies or the existence of structures

on scales larger than 300 million light years. The very existence of galaxies has never been adequately explained by any of the widely accepted theories of cosmology, astrophysics or particle physics.

One could go on, but perhaps the point has been made. The particle physics programme and its offshoots in astrophysics and cosmology are plagued by fundamental gaps, untestable assumptions and *ad hoc* fixes. It is possible that these problems will be overcome gradually in the coming decades, but the contention that we are on the verge of a complete unification of physics, "perhaps close to the final source", is still wishful theoretical whistling against a stiff empirical wind.

ROBERT L. OLDERSHAW

15 West Pelham Road
Shutesbury, Massachusetts 01072, USA

Exploiting authors

SIR—The publication of increasing numbers of multi-author scientific works may be leading to the exploitation of contributing authors. In the past couple of years, I myself have been invited to contribute to four such books, with three US publishers and one European. Each publisher offered to pay about £5 a printed page.

Some very approximate sums indicate what contributors lose by accepting such offers. Consider a book of 400 pages with 10 chapters, each with a different author, the volume costing £50. At £5 a page, each author receives £200; but were 10 per cent royalties (a usual rate for ordinary scientific books of reputable publishers) divided equitably among the authors on a sale of, say, 2,000 copies, each author would receive £1,000. Even if I have overestimated sales by a factor of two, the argument would still be valid.

Five years ago, Watkinson (*Nature* 300, 111; 1982) told us that most of the royalties on multi-authored works went to the editors. This arrangement is absurd as, in my experience, few editors do much editing and, indeed, they seldom earn their 5 per cent, or whatever. Although there must be some honourable exceptions, editors seldom do most of the work.

Young scientists have a special need to publish as much as they can to further their careers. They are delighted to receive any money for an article, which they would usually be glad to have accepted by an academic journal, without remuneration. So they are easy prey for the rapacious publisher. Professional institutes and learned societies might well join together to negotiate minimum rates and other conditions, such as compensation for unduly delayed publication. This would especially benefit those of their members who are inexperienced or do not

have the clout to deal with commercially minded enterprises. The publishers themselves should appreciate that they would also benefit, because appropriate royalties would give authors a stronger motive to complete contributions on time, so that a book would be more up to date when published and, hence, more worthwhile buying.

JAMES A. BARNETT

School of Biological Sciences,
University of East Anglia,
Norwich NR4 7TJ, UK

An artist's eye

SIR—As a corollary of recent correspondence (*Nature* 331, 10; 1988) about eye defects inferred from the paintings of El Greco, what are we to make of Lowry's matchstick men? Can van Gogh's vibrant use of colour be attributed to X-linked colour blindness, or Braque's cubism to possession of a compound eye by goldschmidtian macromutation? The preceding letter on the same page was headed: "God does not need science. . ." — neither does the art critic.

GRAHAM WALLIS

117, Redfern Street,
Subiaco, WA 6008, Australia

SIR—The theory that El Greco had astigmatism was floated in the *Paris Medical Chronicle* in 1913; and it was taken up vigorously in the following decades by (largely German) oculists, who ingeniously countered — as did A.C. Dornhorst (*Nature* 329, 758; 1987) — the obvious objections that had been raised. These were discussed in a book I wrote 18 years ago (*The World Through Blunted Sight*, Thames & Hudson, London, 1970), concluding that in the astigmatic artist "some such influence" was "just conceivable", because, as Nicholas Evans pointed out (*Nature* 331, 10; 1988), perceptual aesthetics can never be linked securely to organic anomalies, such as the shape of the eyeball.

When I first read Sir Peter Medawar's passage ridiculing the "nameless ophthalmologist" who surmised that El Greco was astigmatic (which you so applauded in your leading article 329, 472; 1987) I did draw his attention to the words I had actually used — labelling it as "this time-honoured (and indeed least probable) of all the theories propounding an organic influence on the artist's style". I was hoping that, if proved to be that nameless culprit, I had been misquoted to him, and so might be exonerated. His secretary confirmed that he had received my letter, but declared that he was "too unwell to deal with such correspondence".

PATRICK TREVOR-ROPER

3 Park Square West,
Regents Park,
London NW1 4LJ, UK

The Palaeoindian debate

New data suggest that humans occupied Chile as long ago as 33,000 years ago. But will this end controversy about the date of the first human colonization of the Americas?

THE controversy over the age of the first human colonization of the Americas continues: at the latest meeting on the subject (Smithsonian Institution, Washington DC; September 1987) it became clear that there is still no consensus. Many do not believe humans lived in the Americas until 11,500 years ago, when sites with fluted spear points are abundant and widespread from Alaska to Patagonia. Others argue—in decreasing order of plausibility—for man's arrival somewhere between 30,000 and 40,000 years ago, 200,000 years ago (the notorious Calico Hills site), or even 3,000,000 years ago (at Toca de Esperança in Brazil). The more extreme claims are generally discounted, and the serious debate concerns just a few sites between 20,000 and 40,000 years ago.

In 1986, when the case for human presence at Boqueirão da Pedra Furada, in Brazil, around 32,000 years ago was published in *Nature*^{1,2}, informal opinions among archaeologists ranged from acceptance to complete disbelief. Now, on page 150 of this issue, Dillehay and Collins present new data from Monte Verde, Chile, suggesting human occupation there about 33,000 years ago. But will the sceptics be convinced? The stratigraphy of the site is unambiguous; the radiocarbon dates conform to geological predictions; and there are possible fire hearths—but the artefact sample is dangerously small (8 tools plus 18 naturally split stones imported from outside the area). It is not a good sample, but individually the stone items are similar to those from the upper level, dated to about 13,000 years ago and associated with man-made structures.

In much of the debate about the first Americans, two very different questions have become confused. One is that of the oldest evidence for human presence in the New World. The answer could, in theory, consist of a fire hearth, a butchered carcass, a change in the pollen spectrum, or a piece of stone transported by human agency. Confusion sets in when this question is mixed with that of what the oldest American artefacts look like. Some archaeologists expect to find standardized tool types, analogous to those of the Old World Palaeolithic or the Clovis period in America, whereas others expect the earliest industries to be of a more opportunistic kind, with any suitable piece of stone serving as a tool.

This division within the profession is as much psychological as archaeological, but

it affects the way in which potentially early stone work is evaluated. Some will settle for nothing less than the presence of 'types', or at least of patterned repetition, which is impossible to recognize in a very small sample; others accept more circumstantial kinds of evidence (use-damage, or deliberate selection of raw materials) as sufficient proof of man's presence.

In the case of Monte Verde, readers must decide for themselves. The authors are cautious, claiming only that the site is "a strong contender" for very early status. On previous form, the likelihood is that Monte Verde will not change any of the demarcation lines.

The fact that, after so many years of endeavour, no site older than 12,000 years ago has won universal acceptance, provokes concern about the nature of archaeological 'proof' and about how sites are validated. It is significant that so few archaeologists have changed sides, in either direction, during the past 20 years. Does this mean that the evidence is not yet good enough? Edmund Leach, a leading anthropologist, has said: "Justification in terms of scientific methodology is in part self-deception, for when the figures turn out wrong the true believer will always shuffle the figures; when contrary evidence turns up he throws doubt upon the credentials of the investigator."

Although archaeology likes to masquerade as a science, its objectives are not fundamentally scientific and it does not behave in the same way as physics or biochemistry³. In the 'real' sciences, validation comes through replication of experimental results, but in archaeology this is impossible. Excavation destroys the evidence; once a site is dug we have only the excavator's word for what was found there. In consequence there is always room for doubt. No excavation can ever satisfy the criterion of 100 per cent proof, nor can excavators anticipate what questions will be asked in 20 years' time.

In practice, what saves archaeology from complete chaos is the fact that there is some duplication. When a series of comparable sites has been tested, a general pattern emerges, becomes the new orthodoxy, and governs expectations. This process of validation does not work for the truly unique site and, unfortunately, all the problematic early American sites belong to this one-off category. They are few in number, set in different environments, and spread over

several millennia. Research on Early Man in the Americas is still in a pre-consensus stage, and opinion could go either way. Validation, in the end, depends on the collective professional verdict, itself based on subjective assessment of the investigator's competence and honesty, and the inherent 'reasonableness' of his results.

If the nature of the evidence is a cause for concern, so too is the make-up of the jury. For historical reasons, the international opinion-formers in Palaeoindian studies are mostly English-speaking and based in North America, where convincing early sites have not been found. Inevitably, there are behind-the-scenes mutterings about sour grapes, nationalism, double standards of proof, and a view that some sites (and some researchers) are required to pass more stringent tests than others. True or not, the fact that many people believe these views is disquieting. It is right that new and controversial ideas should be rigorously examined; it was, after all, this kind of persistent scepticism that disposed of the claims of the Old Crow sites⁶ and gave us the correct (Holocene) age of the California skeletons⁷. Does anyone seriously think that a potentially early site dug by, say, a graduate student from Paraguay and published (in Spanish) in a local journal will receive the same kind of assessment as an excavation directed by an established figure, endorsed by other specialists, then published in an international periodical?

There is no obvious way out of this dilemma. Further research will eventually resolve the Palaeoindian question, but the wider problem of how to validate the one-off site will always be with us. Even if there is no solution, I believe we should admit we have a problem and should try to agree what constitutes an acceptable standard of proof when the 100 per cent level is unattainable. Clarifying standards need not mean lowering them. In the meantime we still have to make interim judgements on the sites we have got. **Warwick Bray**

1. Guidon, N. & Delibrias, G. *Nature* **321**, 769–771 (1986).

2. Bray, W. *Nature* **321**, 76 (1986).

3. Dillehay, T.D. & Collins, M.B. *Nature* **332**, 150–152 (1988).

4. Leach, E. *The Explanation of Culture Change* (ed. Renfrew, C.) 762 (Duckworth, London, 1973).

5. Reece, R. *Inst. Archaeol. Bull.* **24**, 191–203 (University College, London, 1987).

6. Nelson, D.E. *et al. Science* **232**, 749 (1986).

7. Bada, J.L. *Am. Antiquity* **50**, 645 (1985).

Warwick Bray is at the University of London Institute of Archaeology, 31–34 Gordon Square, London WC1H 0PY, UK.

Volcano prediction

Lessons from materials science

Robert I. Tilling

In 1902, three cataclysmic eruptions in central America and the Caribbean claimed more than 36,000 lives: at La Soufrière on St Vincent (7 May), Mont Pelée on Martinique (8 May) and Santa Maria in Guatemala (25 October). Subsequently, volcanology entered its modern phase with the establishment of volcano observatories which, by systematically documenting local changes before, during and after eruptions, determine a volcano's present eruptive behaviour. Such data, combined with geological mapping and dating studies to reconstruct the volcano's past behaviour, provide the basis for forecasting future eruptive activity. Most eruptions are preceded by measurable geophysical or geochemical changes, the most diagnostic of which involve variations in magma-induced seismicity and deformation of the volcano's surface¹⁻³. Even with such data, however, successful forecasts are rare. On page 125 of this issue⁴, Voight suggests an intriguing new way of looking at this vexing problem by drawing a mathematical analogy between the terminal stages of material failure and the precursory behaviour of a volcano as it builds towards an eruption. Although Voight's method seems to work well for the few eruptions that he analyses with hindsight, its true predictive potential has yet to be tested.

The notion that rock failure accompanies eruptions is obvious, but Voight extends this idea by arguing that the eruption processes can be described in terms of the fundamental law that governs the failure of materials: $\dot{\Omega} - A = 0$, where A and a are empirical constants; $\dot{\Omega}$ is an arbitrary quantity (dots refer to differentiation with respect to time), such as ground deformation, seismicity or the rate of gas emission. From this general law, Voight derives several specific equations and rate-time plots that can be used to predict rock failure and the attendant eruption. His treatment allows for a time delay between rock failure and eruption. He applies the predictive method to past eruptions (Mount St Helens in the United States, 1981 and 1982; Bezymianny in Kamchatka in the Soviet Union, 1960) and a serious rockslide (Mount Toc in Italy, 1963). In a novel departure from conventional analysis of monitoring data, Voight introduces the concept of inverse-rate-time plots, in which, especially if the inverse-rate curve is nearly linear, the time of rock failure can be estimated by graphical extrapolation of the curve to the intercept on the time abscissa. The modelled inverse-rate curves in some of

the examples considered approach linearity (or $\alpha \approx 2$).

An eruption prediction is a "comparatively precise statement of the time, place, and ideally, the nature and size of impending activity"⁵. Forecasts are less specific, and can be either long-term (made one

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Fig. 1 Mount St Helens in Washington, United States, erupting on 18 May 1980.

year or more in advance) or short-term (hours to months in advance). Both forecasts and predictions generally allow a time window, typically ranging from days to years. A notably successful long-term forecast was made in 1975 by Crandell *et al.*, who stated⁶ that Mount St Helens, which had been dormant since 1857, would be the one volcano in the United States (excluding Hawaii and Alaska) most likely to reawaken and erupt "perhaps before the end of this century". Five years later, the volcano erupted (Fig. 1).

Volcanologists have made successful short-term predictions at some well-monitored volcanoes, most notably in the United States and the Soviet Union^{3,7}. For example, at Mount St Helens, the staff of the Cascades Volcano Observatory since June 1980 have successfully predicted most of the dome-building eruptions, some as much as 3 weeks in advance⁸. For Kilauea volcano, Hawaii, some short-term eruption predictions have even pinpointed the actual site of the eruptive outbreak, so that the affected areas could be evacuated safely.

Most short-term predictions use data from several monitoring techniques rather than rely on any single method. The time

window is estimated largely by empirical methods, using the timing of the onset of anomalous behaviour, rates of changes and comparison with previous events. Voight's method represents a significant refinement in the interpretation of monitoring data, emphasizing the mathematical evaluation of critical rate-change-time relationships.

Voight makes a persuasive case that his approach, with an "experienced interpreter", could improve eruption prediction by recognizing eruption windows sooner than other methods, and by substantially narrowing their width. But his suggestion that often the inverse rate is nearly linear with time ($\alpha \approx 2$) remains to be tested. Also, the examples he analysed include only non-explosive extrusions of viscous lava at Mount St Helens and Bezymianny. The method still needs to be tested at well-monitored volcanoes that erupt fluid basaltic lavas, such as those in Hawaii and Reunion Island. With the possible exception⁷ of the 1975-76 Great Tolbachik Fissure Eruption (in Kamchatka), no major explosive eruptions have been successfully predicted. Two avalanche-triggered eruptions (Mount St Helens, May 1980, and Bezymianny, March 1956), however, were "inevitable" according to a modified form of Voight's technique, although the calculated date has only order-of-magnitude accuracy.

An important unsolved problem is the so-called failed or arrested eruption in which a volcano undergoes the precursory changes typical of impending eruptions but culminates merely with the subsurface intrusion of magma or the end of unrest. The 1983-85 volcanic crisis at Rabaul caldera, Papua New Guinea⁸, provides an excellent example. After the seismic swarm starting in 1983 (Fig. 2) and

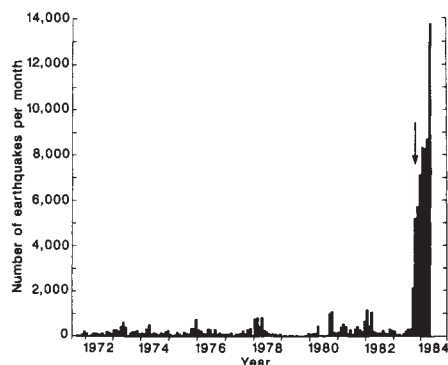


Fig. 2 Seismicity beneath Rabaul caldera in Papua New Guinea from July 1971 to April 1984 (after ref. 8). These data and geodetic measurements prompted volcanologists to forecast a possible eruption within a few months and government officials to issue a stage-2 alert on 29 October 1983 (arrow). After peaking in April 1984, however, the seismicity and caldera deformation declined, returning to 1982 levels by mid-1985. The stage-2 alert was lifted on 22 November, 1984. Could the prediction method of Voight⁴ have done better? Probably not in this case.

accelerating ground-displacement rates, geologists at the Rabaul Volcanological Observatory anticipated a possible eruption within a few months and advised government officials to issue a 'stage-2 alert'. No eruption occurred. In this and probably similar cases, Voight's mathematical analysis of the seismic-energy release and ground-displacement data would have resulted in a more refined, but still false, forecast.

Voight states⁴ that predictions based on fundamental laws "are likely to be more reliable than purely probabilistic attempts based on pattern recognition". Yet, the few examples at only two well-monitored volcanoes analysed by Voight are for eruptions that were successfully predicted by less rigorous techniques^{5,9}. Although Voight's method could improve eruption forecasting, its use will be limited until it has been tested for more eruptive styles and volcano types. Until then, volcanologists will rely on recognizing patterns,

because the quantitative data needed to apply Voight's method are inadequate or lacking for most volcanoes. Indeed, if the science (art?) of eruption forecasting is to improve, a global programme is sorely needed to establish baseline monitoring networks at active and potentially active volcanoes. One such programme, recently proposed¹⁰, could represent a start in this direction. □

1. Tazieff, H. & Sabroux, J.-C. (eds) *Forecasting Volcanic Events* (Elsevier, Amsterdam, 1983).
2. Souther, J.G., Tilling, R.I. & Punongbayan, R.S. *Episodes* 7, 10-17 (1984).
3. Decker, R.W. *A. Rev. Earth planet. Sci.* 14, 267 (1986).
4. Voight, B. *Nature* 332, 125-130 (1988).
5. Swanson, D.A. *et al. J. Geodynamics* 3, 397-423 (1985).
6. Crandell, D.R. *et al. Science* 187, 438-441 (1975).
7. Tokarev, P.I. *Bull. volcan.* 41, 251-258 (1978).
8. McKee, C.O. *et al. J. Volc. geotherm. Res.* 23, 195 (1985).
9. Tokarev, P.I. *Bull. volcan.* 26, 379-386 (1963).
10. Housner, G.W. (ed.) *Confronting Natural Disasters: An International Decade for Natural Hazard Reduction* (National Academy, Washington DC, 1987).

Robert I. Tilling is at the US Geological Survey, 345 Middlefield Road, Menlo Park, California 94025, USA.

Malaria

Effective vaccine for humans?

Louis H. Miller

THE outlook for malaria is dismal. In a recent study in Africa, it was estimated that 25 per cent of the deaths in children 1-4 years of age were caused by malaria¹. With the spread of chloroquine resistance, the death rate may increase. Insecticides and other mosquito-control methods are largely ineffective in Africa. It is for these reasons that an effective vaccine against falciparum malaria would be one of the most important steps towards improved health in the tropical world, particularly for children in Africa. The report by Patarroyo *et al.* on page 158 of this issue² describes the first human trial of a malaria vaccine against the asexual erythrocyte parasite of *Plasmodium falciparum*.

Vaccines will probably contain antigens specific for three stages of the parasite: sporozoites, asexual erythrocyte parasites and sexual forms³. Immunity induced by these vaccines would, in general, be stage-specific. Thus, immunity to sporozoites, the stage of the parasite that is injected by mosquitoes, would not protect against the asexual erythrocyte stage, which causes disease and death. If a few sporozoites evade sporozoite immunity, the resulting infection could be fully virulent. On the other hand, a vaccine against the asexual erythrocyte parasite, even if only partially effective, would reduce the severity of disease.

Unlike some viruses that induce protection after a single infection, malaria parasites can reinfect humans repeatedly over their entire lifetime. The parasite evades immune attack through antigenic

variation, mutations, the poor immunogenicity of surface antigens and the modulation of the immune response (suppression). Despite this, people in endemic areas do eventually develop a degree of immunity to malaria that permits survival in the face of repeated infections. Infection in the immune adult produces few symptoms.

The challenge facing development of a subunit vaccine against asexual erythrocyte parasites is formidable. How does one determine which of the many parasite proteins (and correspondingly larger number of peptides) should be included? One approach is to identify the proteins recognized by immune sera⁴, making the assumption that the best antigens for vaccines are recognized by sera from immune individuals and that the best type of immunity is antibody-mediated. This

approach may not identify the best immunogens for a T-cell-mediated immunity, although an antigen identified by antibody may also contain effective immunogens for T-cell-dependent immunity.

Another approach is to study the T-cell response in infected populations to antigens identified by other means (such as antibody)⁵. It still remains to be seen which of these proteins or peptides that is recognized by antibody or induces proliferation of T cells also induces protection. The limited supply of New World monkeys prevents the testing of more than a few of these proteins.

A third approach is the identification of proteins on the merozoite surface, the stage of the parasite that invades erythrocytes, or on the surface of infected erythrocytes. In addition, there are attempts to identify molecules involved in critical events such as erythrocyte invasion or cytoadherence of parasitized erythrocytes to vascular endothelium. The assumption is that antibodies to these proteins will inhibit invasion of erythrocytes or promote clearance of infected erythrocytes. Unfortunately, the main immunogen on the erythrocyte surface undergoes antigenic variation^{6,7}, and the merozoites of any particular malaria may have alternative pathways for invading erythrocytes⁸. A variant of this approach is the continuing search for monoclonal antibodies that block invasion⁹. The proteins recognized by these blocking antibodies could then become 'vaccine candidates'.

Patarroyo *et al.* used none of these approaches but directly identified antigens of asexual stages of the parasite that would protect *Aotus trivirgatus*, a New World monkey (see figure), against fatal infection with *P. falciparum*. They had previously purified detergent-solubilized proteins from mature asexual erythrocyte parasites on SDS-polyacrylimide gels and had immunized monkeys with the proteins from different regions of the gel^{10,11}, and determined the amino-terminal sequence of the predominant protein in the two protective fractions. These two peptides, when combined with a previously described



Aotus monkeys are collected by Indians in the Amazonian basin. With a group of monkeys inside the hollow tree trunks that they live in, the Indians first seal off all the exit holes. They then reopen one hole and, as the enraged animals emerge, try to catch them by the tail.

It can take a week to catch a single animal. Captured monkeys are taken to Leticia at the southernmost tip of Colombia, where the malaria trials are carried out. Because the Aotus monkey is a protected species, those that succumb to malaria have to be cured with drugs. They are then passed on to a breeding colony elsewhere in Colombia and, sometimes, released back into the wild. (Photo courtesy of M. Patarroyo.) □

one, led to a high level of protection in monkeys. In their work on humans², Patarroyo *et al.* synthesized a 45-amino-acid peptide (SPf-66) containing these three peptides and polymerized it into a high-molecular-weight protein. Patarroyo *et al.* find that this polymer in alum induces protective immunity in human volunteers. It is worrying that one of the peptides used in this vaccine has two different amino acids in one of four isolates of *P. falciparum* sequenced¹², because vaccination of a population could rapidly select variants that are unaffected by immunity induced to the vaccine.

When it comes to testing a human vaccine against *P. falciparum*, there is a particular difficulty of evaluation. The asexual erythrocyte parasite develops in 48 hours from a young parasite, the ring form, to a mature parasite that can infect ten other erythrocytes. Because mature parasitized erythrocytes are sequestered along vascular endothelium, only erythrocytes containing ring forms are seen in the peripheral blood. After the parasitaemia reaches 0.5 per cent ring forms, for example, it decreases as the parasitized erythrocytes are sequestered or as a result of immunity. There is no way of knowing if the parasitaemia will jump, 48 hours later, to 5 per cent, a dangerously high level, or if immunity has controlled parasite development. Thus, no study of the efficacy of an asexual erythrocyte vaccine is without some risk.

The authors, aware of this dilemma, chose to treat patients at 0.5 per cent parasitaemia. But in the case of four volunteers, treatment was not initiated promptly, resulting in one of the volunteers (J.D.) reaching a parasitaemia of 4.26 per cent. The failure strictly to follow the protocol placed some volunteers at excessive risk. Had the level of parasitaemia for initiating therapy been set lower, for example, 0.25 per cent as I would have preferred, only one volunteer (W.B.) would have been protected. This illuminates a problem that will be faced in testing vaccines to control asexual parasitaemia. This contrasts with trials of sporozoite vaccines where volunteers are at minimal risk, as they are treated as soon as a parasite is observed in the blood.

What is the next step now that the first human trial of an asexual erythrocyte vaccine is completed? The top priority is the need to confirm the monkey trials in other laboratories and with other isolates of *P. falciparum*. When confirmed, the trials should be extended to include other proteins or peptides (such as the recombinant protein of RESA/Pf155)¹³ to be tested in monkeys. Better adjuvants or delivery by a living organism (such as BCG or attenuated *Salmonella*) may be required. There still remains the indeterminacy of when a vaccine will be available for developing countries.

The main concern now is the selection of mutant parasites that are resistant to a vaccine. Inclusion of a vaccine against the sexual stages of the parasite can block the transmission of mutant parasites in vaccinated people. Because of the indeterminacy, countries in the developing world must continue to devise methods within their means to control malaria and to limit disease, and the research base must remain broad, including research on chemotherapy and on the mosquito vector. DDT was said to have eliminated malariologists in the 1950s because it appeared so promising. Why did scientists at that time leave malaria research for other fields although Africa was largely untouched by insecticides? This type of short-sightedness should not now dominate research priorities. □

1. Greenwood, B. M. *et al.* *Trans. R. Soc. trop. Med. Hyg.* **81**, 478–486 (1987).
2. Patarroyo, M. E. *et al.* *Nature* **332**, 158–161 (1988).
3. Miller, L. H. *et al.* *Science* **234**, 1349–1356 (1986).
4. Kemp, D. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 3787–3791 (1983).
5. Good, M. F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 1199–1203 (1988).
6. Brown, K. N. & Brown, I. N. *Nature* **208**, 1286 (1965).
7. Hommel, M., David, P. H. & Oligino, L. D. *J. exp. Med.* **157**, 1137–1148 (1983).
8. Hadley, T. J. *et al.* *J. clin. Invest.* **80**, 1190–1193 (1987).
9. Klotz, F. W., Hudson, D. & Miller, L. H. *J. exp. Med.* **165**, 359–367 (1987).
10. Patarroyo, M. E. *et al.* *Nature* **328**, 629–632 (1987).
11. Patarroyo, M. E. *et al.* in *Modern Approaches to New Vaccines: Prevention of AIDS and other Viral, Bacterial, and Parasitic Diseases* (eds Chanock, R. M. *et al.*) 117–124 (Cold Spring Harbor, New York, 1987).
12. Mackay, M. *et al.* *EMBO J.* **4**, 3823–3829 (1985).
13. Collins, W. E. *et al.* *Nature* **323**, 259–262 (1986).

Louis H. Miller is in the Malarial Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

I. I. Rabi (1898–1988)

FOR the last decade of his life, I. I. Rabi was the acknowledged dean of physics in the United States. He died on 11 January and with his passing goes a man of vision — the moral conscience of physics since the Second World War.

After completing his PhD at Columbia University in 1926, Rabi went to Europe to learn the new physics, studying briefly with Schrödinger, Sommerfeld, Bohr, Pauli and Heisenberg. But his work with Otto Stern during 1928 determined the course of his physical research. He returned to the United States and Columbia University in 1929. After two years of uninspired attempts to continue his theoretical work on solids, Rabi decided to go back to molecular-beam studies.

In 1931, three years after Dirac's seminal work on the quantum theory of the electron, the basic physics of the atom seemed well understood. It was the atomic nucleus that provided the questions at the frontier of physics. With the Breit-Rabi theory, developed in 1931, molecular-beam methods could be used to study the magnetic properties of the atomic nucleus — even though nuclear magnetic moments are 2,000 times smaller than those of the atom.

Rabi started with the basic Stern-Gerlach method and, through the 1930s, he added successive magnetic fields, each of which worked its influence on the tiny nuclear magnetic moments of the passing beam particles. By 1938, he had developed the magnetic resonance method, the precursor to both nuclear magnetic resonance in bulk matter and magnetic imaging. Many atoms coursed through his molecular-beam apparatus, but hydrogen was the one that captured Rabi's imagination. "Here you have a system you could understand", said Rabi. "There were no complications. Anything I couldn't understand was because there was something to be discovered."

Rabi's measured values of the proton's magnetic moment reduced the experimental uncertainties from 10 to 5 to 0.7 per cent. As a finale, he discovered, with J.M.B. Kellogg, N.F. Ramsey and J. Zacharias, the quadrupole moment of the deuteron. In 1944 Rabi won the Nobel prize for physics.

The Second World War brought a sudden halt to Rabi's research. With the cavity magnetron invented at the University of Birmingham as the centrepiece, the radar project began at the Massachusetts Institute of Technology where Rabi was the associate director of the radiation laboratory. Later, along with Bohr, Rabi served as senior advisor to J. R. Oppenheimer on the Manhattan Project.

After the war, Rabi became a statesman of science. He believed that science could be used to break through barriers that separated nations and he put this belief to practice. Rabi was instrumental in founding the Brookhaven National Laboratory in the United States and, with that as a model, he proposed to the fifth general assembly of UNESCO, in 1950, the establishment of a European nuclear research laboratory. Eighteen months later, eleven European nations signed the document establishing CERN. In 1955, Rabi saw another of the products of his efforts come to fruition: the first international conference on the peaceful uses of atomic energy, which was held in Geneva in Switzerland.

Rabi's interest in physics remained acute to the very end of his life. Each time I visited him, he wanted to talk about "the subject". Shortly before he died, Rabi said to me, "I'm going for 90". He missed by 6 months, an experimental error he would have loathed.

John S. Rigden

John S. Rigden is Editor of the *American Journal of Physics*, 353 East 45th Street, New York, New York 10017, USA.

Photosynthesis

Signals from the reaction centre

Jim Barber

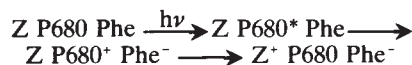
OUR understanding of the molecular mechanisms by which photosynthetic organisms use solar energy to bring about the oxidation of water is being greatly aided by the technique of protein engineering. The latest step using this method, reported independently by two groups^{1,2}, identifies and locates an amino-acid residue in the photosynthetic reaction centre that could be responsible for the transfer of electrons from the water-splitting complex to the primary oxidant P680⁺.

The oxidation of water by photosynthetic organisms results in the release of molecular oxygen into the atmosphere, a process on which all animal life is dependent. The other products of water oxidation are hydrogen ions (H⁺) and electrons (e⁻), of which there are four of each for every molecule of oxygen produced.



The hydrogen ions and electrons thus produced do not combine to form molecular hydrogen but reduce atmospheric carbon dioxide to organic compounds.

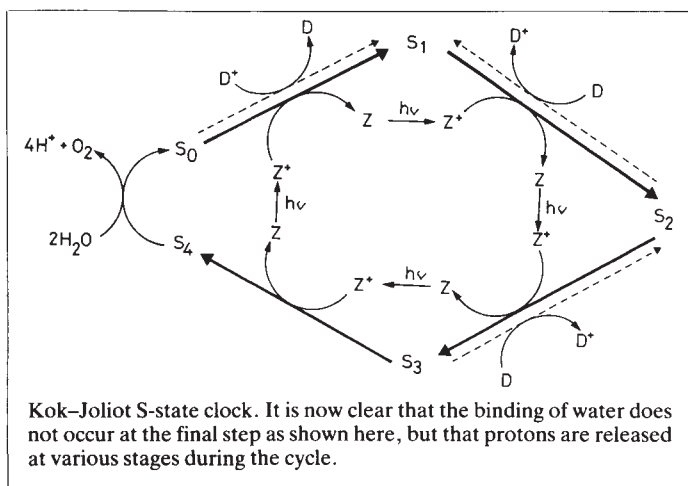
The photons of visible light used to split water are absorbed by a pigment-protein complex in the membrane called photosystem two (PS2). This complex consists of two functional units, one, the reaction centre, which is photochemically active, and the other, the light-harvesting system, which consists of several hundred chlorophyll molecules and acts as an 'antenna' for the reaction centre. Excitations created in the antenna chlorophylls are rapidly transferred to a photochemically active site where very rapid charge separation occurs.



P680 is a form of chlorophyll distinct from those of the antenna system; Phe is a pheophytin molecule; and Z is the primary electron donor to P680⁺. The redox state of Z⁺ must be very high, about +1 V, to act as an effective oxidant of water. Moreover, the charge-transfer process must turn over four times to accumulate the four oxidizing equivalents necessary to produce one oxygen molecule. This requirement for charge storage can be described in terms of the model of Joliot

and Kok (ref. 3; see figure).

The accumulation of charge in the various S-states involves valency changes in at least two, and maybe four, manganese atoms which are bound to the PS2 complex. Until now, the identity of Z has been unknown. Its existence has been monitored by electron paramagnetic resonance (EPR) spectroscopy and is detected with a spectroscopic-splitting value of $g=2.0046$. The EPR signal at this g -value, however, has two kinetically dis-



Kok-Joliot S-state clock. It is now clear that the binding of water does not occur at the final step as shown here, but that protons are released at various stages during the cycle.

ting phases when illumination is stopped: a fast phase characteristic of the decay of Z⁺; and a slow phase indicative of another component, D⁺. The function of this latter oxidant is unknown but it is not directly involved in 'main-line' electron transfer to P680. It does, however, have the capacity in the dark to convert S₀ to S₁ and, in its non-oxidized form, to bring about a dark reduction of S₂ and S₃ (see figure). The overall effect is that S₀ accumulates in the dark⁴. Arguments have been presented to suggest that Z⁺ and D⁺ are radicals of quinone molecules tightly bound to the PS2 complex⁴. Some elegant recent work shows that these arguments are incorrect and that Z and D are almost certainly tyrosine residues within the PS2 reaction-centre polypeptides (see below).

As I reported previously⁵, it is now established that P680 and Phe are bound to a complex consisting of two polypeptides, D1 and D2. Based on amino-acid sequence similarities, the heterodimer which they form is likely to have a very similar structure to that of the L and M subunits of the reaction centre of purple bacteria⁶. As the three-dimensional structure of the purple bacterial reaction centre is known to a resolution of atomic distances^{7,8}, it has been possible to speculate about the significance of the amino-acid

homologies in terms of chromophore binding sites. The purple bacterial reaction centre, however, unlike PS2, cannot bring about the oxidation of water and therefore does not generate EPR signals indicative of Z⁺ and D⁺.

Hints that Z and D could be tyrosine residues on the D1 and D2 polypeptides came from studies using radioactive iodine^{9,10}. This idea has recently been nicely substantiated by Barry and Babcock¹¹. These authors fed deuterated tyrosine to the oxygenic cyanobacterium, *Synechocystis*, under conditions in which the amino acid was not degraded and was incorporated into the protein of the organism. They found that the normal 20 Gauss linewidth of the $g = 2.0046$ EPR signal is narrowed to 7 Gauss in cultures fed with deuterated tyrosine. Such a narrowing is consistent with the notion that the EPR signals indicating Z⁺ and D⁺ are indeed caused by tyrosine radicals.

The earlier iodination experiments indicated that Z is on the D1 and D on the D2, suggesting that there is a tyrosine conserved between these two polypeptides. In fact, bearing in mind the probability that the D1/D2 heterodimer is likely to have structural features similar to the L/M subunits of the bacterial system⁵, tyrosine residues at positions 160 or 161 seemed likely candidates for Z

and D. Indeed, in support of this notion, no such tyrosines are conserved in this symmetrical way on the L and M polypeptides.

This possible identity for D has been elegantly demonstrated very recently by two independent groups, who have applied the technique of site-directed mutagenesis. Using a transformable strain of *Synechocystis*, Debus and co-workers at Michigan State University¹ specifically mutated tyrosine 160 on the D2 polypeptide to a phenylalanine. They find that the resultant mutant lacks the EPR signal indicative of D⁺. Precisely the same result was obtained independently by Vermaas and colleagues at Arizona State University². In addition to mutating Tyr 160 on D2 to Phe, these workers also mutated the flanking methionine residue at 159 to an arginine. This latter mutation, unlike the conversion of Tyr 160 to Phe, does not result in the loss of the EPR signal indicative of D⁺.

Both groups note that the Tyr 160-Phe mutant could grow photosynthetically, but at a significantly slower rate than for the wild-type organism. Vermaas and colleagues show that this slower rate is not caused by changes in the kinetics of electron donation from Z to P680 or to a decrease in the density of PS2 reaction

centres. Therefore it seems that D⁺ is not an absolute requirement for photosynthetic activity but does play some role in optimizing metabolic activity.

The very important conclusion which can be drawn from this work, given that D⁺ is a tyrosine radical at position 160 on the D2 polypeptide of *Synechocystis*, is that Z is almost certainly the symmetrically related tyrosine on the D1 polypeptide. By analogy with the structure of the L and M subunits, it seems that the tyrosines are probably located in the third membrane helices of the two reaction-centre polypeptides and are at distances from the membrane surface equivalent to that expected for P680, assuming histidines 198 on D1 and D2 form ligands with the chlorophylls which constitute this primary electron donor.

It is particularly surprising that the similarity of the environments of two tyrosines of D1 and D2 can give rise to dramatically different redox kinetics of Z and D. This, therefore, is another striking example of

major functional differences between structurally symmetrical components within photosynthetic reaction centres that has been so clearly shown for bacteriochlorophyll and bacteriopheophytin of the purple bacterial system^{7,8}. □

1. Debus, R.J., Barry, B.A., Babcock, G.T. & McIntosh, L. *Proc. natn. Acad. Sci. U.S.A.* **85**, 427–430 (1988).
2. Vermaas, W.F.J., Rutherford, A.W. & Hansson, O. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
3. Joliot, P. & Kok, B. in *Bioenergetics of Photosynthesis* (ed. Govindjee) 387–412 (Academic, New York, 1975).
4. Babcock, G.T. in *Photosynthesis* (ed. Amez, J.) 125–152 (Elsevier, Amsterdam, 1987).
5. Barber, J. *Trends biochem. sci.* **12**, 321–326 (1987).
6. Evans, M.C.W. *Nature* **327**, 284–285 (1987).
7. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *Nature* **318**, 618–624 (1985).
8. Allen, J.P., Feher, G., Yeates, T.O., Komiya, H. & Rees, D.C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5730–5734 (1987).
9. Takahashi, Y., Takahashi, M. & Satoh, K. *FEBS Lett.* **208**, 347–351 (1986).
10. Ikeuchi, M. & Inoue, Y. *FEBS Lett.* **210**, 71–76 (1987).
11. Barry, B.A. & Babcock G.T. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7099–7103 (1987).

Jim Barber is Professor of Plant Physiology and Head of the AFRC Photosynthesis Research Group, Imperial College of Science and Technology, London SW7 2BB, UK.

Materials science

Strategies to defeat brittleness

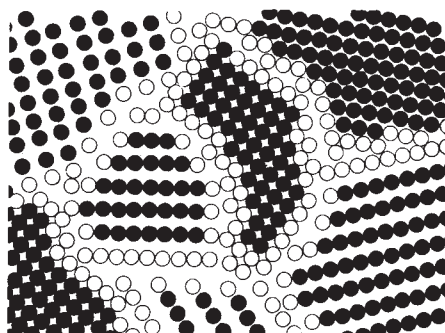
Robert W. Cahn

THE highest attainable strength is found in materials composed of light atoms linked into a network by strong ionic or covalent bonds — in other words, ceramics. These also have particularly high melting-points and so would be apt for engineering use at both ambient and high temperatures were it not for their universal brittleness, which is a synonym for the inability to deform plastically and thus to absorb mechanical or thermal shocks. No ultrabrittle material could ever be trusted, for instance, in an aero engine. Two recent papers^{1,2} present radically new ways of overcoming brittleness in hard materials. Gleiter *et al.* show¹ that reducing the grain size of ceramics can allow them to flow plastically. And on page 139 of this issue², Brookes *et al.* show that even diamond, if pressurized in the right way, can deform without fracturing.

During the past 15 years particularly, attempts have been made to overcome the handicap of brittleness by improvements in processing and composition³. Macro-defect-free cement⁴, chemically modified silicon nitrides⁵ and transformation-toughened zirconia^{6,7} represent successful improvements based on these two strategies. The last of these was originally proposed⁸ with the provocative title "Ceramic steel?", because doped zirconia, like high-tensile steel, depends on a partial phase transformation to temper strength with shock resistance: an advancing crack triggers a local transformation which hinders further propagation of the crack.

None of these improvements has gone far enough. Jack⁹ described with feeling the technological, economic and organizational problems that dimmed the promise of silicon nitride, which "has now been the ceramic of the decade for three decades". The ceramic engine, outside Japan at least, still seems a long way from practical realization. There is scope for radically new approaches.

One such approach, recently reported by Gleiter and co-workers in *Nature*¹, depends on the creation of what is virtually a new state of matter, intermediate between a polycrystal and a glass — the 'nanocrystal'. The structure of a nano-



Structure of a nanocrystalline metal. Solid circles, atoms arranged in crystalline structures; open circles, loosely packed intergranular atoms that give the material its high self-diffusivity and plasticity. Under pressure, atoms on the compressed side of a grain rapidly self-diffuse through the intergranular circuit to the stretched side. (Courtesy of H. Gleiter.)

crystalline metal, made by evaporation or sputtering onto a cold substrate followed by compaction of the fine powder so produced, is shown schematically in the figure. The crystal grains are so small that almost half the atoms in the metal occupy intergranular sites; examination by X-ray diffraction of nanocrystalline gold shows⁹ that the intergranular material is gas-like rather than glass-like — there is no short-range order. Mössbauer spectroscopy of nanocrystalline iron over a range of temperatures shows¹⁰ that the intergranular matter has a lower Debye temperature (or elastic constant), and thus has a lower cohesion, than do the grains.

The solubility of hydrogen as a function of its concentration is a measure of the prevalence of microcavities, which are hydrogen traps; such solubility measurements on nanocrystalline palladium show¹¹ that the intergranular 'phase' is indeed full of cavity-traps. This in turn implies, as has been observed¹¹, that solute diffusion of small atoms along the intergranular short-circuits is extremely concentration-dependent and can be higher or lower than in the crystal. Self-diffusion, however, is not affected by traps and is spectacularly enhanced. Self-diffusion at 20–120 °C in nanocrystalline copper is enhanced by no less than 19 orders of magnitude compared with an ordinary polycrystal¹².

Gleiter and co-workers observe¹ easy plasticity, at 80–180 °C, of nanocrystalline TiO₂ (made by oxidation of nanocrystalline titanium) and CaF₂ with grain diameters of about 10 nm — smaller by a factor of about 1,000 than normal grains. This can be interpreted in terms of diffusion creep. In this process, the cations and anions diffuse from the compressed to the stretched side of each grain, round an intergranular circuit, thereby changing the shape of the grain. The authors estimate that the small grain size enhances the strain rate by a factor of 10⁹, and the large self-diffusivity by a further factor of 1,000. In effect, these materials behave superplastically¹³, although the superplastic strain rates that are shown to be feasible are very much higher than those normally encountered in the industrial superplastic forming of aluminium alloys.

Of course, a superplastic ceramic will be weak in a critical range of temperatures: but this can in principle be counteracted by heating the ceramic to make the grains grow slightly: the larger the grains, the higher the critical temperature range. As the authors point out, the core and surface of a ceramic block can be treated (perhaps by microwave heating) to have different grain sizes and plastic properties. The possibilities are intriguing.

Gleiter and colleagues found it necessary to deform their ceramics by compressing them in contact with a platen

made of a soft metal, aluminium, backed by a harder metal; presumably, contact with a hard platen causes intragranular cracking before the diffusion creep can begin. This soft-plate technique is similar to that used in the very different study of Brookes *et al.* reported in this issue². These authors achieve plastic deformation in a monocrystal of diamond, the hardest and most brittle of solids. Following earlier work with MgO and TiC, they reason that a conical indenter pressed on a smooth diamond surface under constant load will — so long as it is slightly softer than the diamond — slowly be blunted until the distributed contact stress has fallen sufficiently for deformation to cease. If the hardness of the indenter is chosen correctly, then the stress in the contact zone is too low to crack the diamond but high enough to deform it plastically. For this to be feasible, the temperature must be very high. An indenter of cubic boron nitride (second only to diamond in hardness) at 1,000 °C neatly does the trick. The diamond deforms by clearly visible multiple dislocation glide. This temperature is 500 °C lower than the hitherto-accepted transition from brittleness to slight ductility, as determined by three-point bending¹⁴.

It would be interesting to know how the grain size of polycrystalline diamond affects its plastic deformability; indeed, it is conceivable that a sufficiently fine grain size could permit a measure of diffusion creep, not linked to the motion of dislocations. In view of the extensive recent work on the formation of microcrystalline-diamond surface coatings by vapour deposition (see ref. 15), such an investigation could be of great practical value. □

1. Karch, J., Birringer, R. & Gleiter, H. *Nature* **330**, 556–558 (1987).
2. Brookes, C.A., Howes, V.R. & Parry, A.R. *Nature* **332**, 139–141 (1988).
3. Brook, R.J. in *Encyclopedia of Materials Science and Engineering* (ed. Bever, M.B.) 98–105 (Pergamon and MIT, 1986).
4. Birchall, J.D. *Phil. Trans. R. Soc. A* **310**, 139–153 (1983).
5. Jack, K.H. in *High-Technology Ceramics, Past, Present and Future; Ceramics and Civilization* Vol. III (ed. Kingery, W.D.) 259–288 (Am. Ceramic Soc., Westerville, 1986).
6. Claussen, N. & Heuer, A.H. in *Encyclopedia of Materials Science and Engineering* (ed. Bever, M.B.) 5129–5132 (Pergamon and MIT, 1986).
7. Lange, F.F. *J. mater. Sci.* **17**, 225–263 (1982).
8. Garvie, R.C., Hannink, R.H. & Pascoe, R.T. *Nature* **258**, 703–704 (1975).
9. Zhu, X., Birringer, R., Herr, U. & Gleiter, H. *Phys. Rev. B* **35**, 9085–9090 (1987).
10. Herr, U., Jing, J., Birneger, R., Gonser, U. & Gleiter, H. *J. appl. Phys.* **50**, 472–474 (1987).
11. Mütschele, T. & Kirchheim, R. *Scripta Metall.* **21**, 135–140 (1987).
12. Horvath, J., Birringer, R. & Gleiter, H. *Solid St. Commun.* **62**, 319–322 (1987).
13. Cahn, R.W. & Hazzledine, P. in *Encyclopedia of Materials Science and Engineering* (ed. Bever, M.B.) 4686–4790 (Pergamon and MIT, 1986).
14. Evans, T. & Wild, R.K. *Phil. Mag.* **12**, 479–486 (1965).
15. Messier, R., Spear, K.E., Badzian, A.R. & Roy, R. *J. Metals*, N.Y. **39** (9), 8–11 (1987).

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

Myth, science and art of Greek peonies

At an exhibition at the Natural History Museum, Cromwell Road, London SW7, UK until 17 April the peony is used to trace the development of scientific documentation and botanical illustration. Peonies grow wild on the mountains and hills of Greece (the example shown here is the species *Paeonia parnassica* from Mount Parnassos), and were used 2,000 years ago for medical purposes, being associated with the mythical physician Paeon. The exhibition contains twelve lithographs made from paintings by Niki Goulondris. The lithographer, Takis Katsoulis, used 12–16 zinc plates for each illustration, one for each colour variation. The lithographs were made by printing each plate on a stone or metal press and then superimposing by hand. □

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Courtesy of Niki Goulondris

Developmental biology

Growth factors in amphibian cell differentiation

Hugh Woodland and Liz Jones

THE molecules that control development are being unveiled at an accelerating rate. What is more striking than anything else is that so many fall into a select band of molecular families. These families, identified by their first members, are growth factors and their receptors, oncogenes and DNA-binding proteins with either zinc fingers or homoeo domains. Very recent examples are *wingless*, a segmentation gene of the fruitfly *Drosophila*, related to *int-1*, a mouse oncogene activated by mammary tumour virus¹; and *dorsal*, which is involved in formation of the fly's dorso-ventral axis, and has extensive homology with *v-rel*, an avian oncogene². Growth factors have an even longer pedigree in this regard. For example, epidermal growth factor (EGF) relatives have a role in cell-fate switching in the nematode reproductive system (*lin-12*)³ and between epidermis and nervous system in *Drosophila* (*notch*)⁴.

Two papers, one by Kimelman and Kirschner⁵ and the other by Weeks and Melton⁶, published in a recent issue of *Cell*, suggest that growth factors or their relatives have a role in mesoderm

formation and the establishment of the dorso-ventral axis in Amphibia.

Dorso-ventral polarity

When the amphibian egg is laid it has an obvious pigmented animal pole and a yolky vegetal pole (Fig. 1a). The animal cells develop into ectoderm and the vegetal cells into endoderm. These cells also interact in the blastula, the vegetal cells inducing neighbouring animal cells to become mesoderm. The animal and vegetal cells are the first types to appear and presumably their different properties are established by substances ('determinants') localized in the egg. Without polarization phenomena, triggered by the entry of the sperm, the embryo would develop into a tube of radially symmetrical ventral tissues.

The effect of the sperm is to cause a polarized contraction of the surface regions of the egg⁷. In turn, this somehow makes the vegetal cells on the dorsal side generate a dorsal mesoderm induction signal (B, Fig. 1b) rather than, or together with, the basal ventral signal (A, Fig. 1b). The dorsal mesoderm subsequently organizes many later developmental

processes, for example, inducing contiguous ectoderm cells to form the nervous system. The most dorsal mesoderm also emits a dorsalizing signal which adjusts the detailed dorso-ventral patterning of the mesoderm (signal C, Fig 1c). This is seen in grafts when dorsal mesoderm is juxtaposed with ventral mesoderm; the ventral mesoderm takes on more dorsal development⁸. In whole embryos, such a graft produces the classic Spemann–Mangold siamese twins, with two complete sets of dorsal structures (brains, spinal cords, notochords and so on).

Until recently, nothing was known about the molecular nature of these stimuli, nor even if A, B and C represent different kinds of molecules, rather than different amounts of the same substance. C is probably different from A and B, however, as mesoderm does not go on inducing more mesoderm in naïve animal cells.

Growth factors

The first clue that these inductive agents might relate to growth factors came from the work of Slack and collaborators⁹, who tested every available growth factor on animal cap cells and found that two of them induce the appearance of mesoderm: fibroblast-derived growth factor (FGF) and embryonal carcinoma-derived growth factor. Both are in a family of growth factors which bind to heparin and, as mentioned above, relatives have precise morphogenetic effects in switching cell fate in nematodes and insects. The effectiveness of FGF in inducing mesoderm has now been confirmed in the new work of Kimelman and Kirschner⁵. However, the interpretation of these authors relies solely on a single molecular marker, embryonic muscle actin messenger (m) RNA. This marker identifies only developing muscle and therefore detects solely dorsal mesoderm (though not the most dorsal mesoderm, the notochord; Fig 1d). Kimelman and Kirschner find that FGF will induce only small amounts of muscle and hence conclude that FGF is a weak mesoderm inducer. However, the results of Slack *et al.*⁹ show that FGF actually induces large amounts of mesoderm, but that it is predominantly ventral in nature. This immediately suggests that FGF could be related to signal A in Fig 1.

Kimelman and Kirschner's more important contribution is to show that a second mammalian growth factor, transforming growth factor-beta (TGF- β), acts synergistically with FGF to increase the amount of muscle actin mRNA in animal cap cells. It is not clear that it actually increases the amount of mesoderm, but it does switch its nature to being much more

dorsal than with FGF alone (though never so dorsal as to contain notochord, except in very rare cases; J. Slack, personal communication). Kimelman and Kirschner suggest that a molecule related to TGF- β might be part of signal B in Fig 1.

As it stands, these hypotheses are relevant only if such molecules actually exist in the embryo. It is therefore of the greatest significance that Kimelman and Kirschner were able to isolate, from an

substance, which suppresses the development of the female reproductive system in males. In the carboxy-terminal region, which corresponds to the cleaved active peptide, Vg-1 is 38% identical to TGF- β , about 40% to inhibin, and particularly interestingly, 48% to a *Drosophila* protein, the *decapentaplegic* product. This is encoded by a zygotic gene which is essential for the fly's dorsal development. These observations lead to the working

hypothesis that FGF and Vg-1 cooperate to form the dorso-ventrally polarized mesoderm induction signal: FGF is signal A, and Vg-1 is signal B.

It is essential to be cautious in accepting this hypothesis, useful and exciting though it is. The activity of none of the *Xenopus* molecules has been tested, and this includes Vg-1. We mention above how diverse the effects of the TGF- β family can be. It is conceivable that Vg-1 does not have the function proposed, or at least that other molecules are involved. Our work (unpublished) shows that mesoderm induction in the tail region persists long after Vg-1 mRNA becomes undetectable¹⁰, indicating that other molecules could be involved.

Nevertheless, at least Vg-1 mRNA is in precisely the correct region of the embryo (Fig 2). Initially in oocytes it is found tightly localized to a cortical shell at the vegetal pole. In the egg and cleaving embryo, the localization slowly breaks down and the mRNA spreads through the yolky, vegetal region that later forms the endoderm and throughout which the mesoderm-induction signal exists^{8,12}. The precise localization of the Vg-1 mRNA in the dorsoventral axis has not yet been determined, but it does not seem to be grossly concentrated to the future dorsal side. Thus, if Vg-1 does what is postulated, dorsalization probably represents either localized translation of the mRNA, or post-translational activation of the protein. Translational control phenomena are known to be very important in these early embryos¹³.

Mesoderm induction probably starts at the 32/64-cell stage. Interestingly this is just the stage at which injecting Li^+ into ventral signalling vegetal cells can generate a new dorsal signal¹⁴. We do not know how Li^+ works, but one of its known effects is on the inositol phosphate/diacylglycerol second-messenger system, which may therefore be involved in the post-transcriptional activation of stored mRNAs, or their protein products.

Although we are now entering the realms of speculation, it now looks very likely that mRNAs for growth factors are examples of the localized determinants in

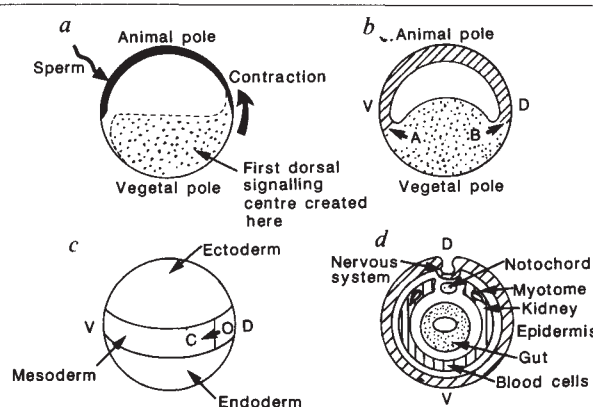


Fig. 1 Early events in forming the dorsoventral axis of an amphibian embryo. *a*, Section of a fertilized egg. A contraction of the cortex of the egg is far stronger on the side opposite sperm entry. This dorsalizes the vegetal region of this side, so that from cleavage stages onwards (*b*) the cells derived from it generate a dorsal mesoderm induction signal, B. The cells of the other side give a ventral mesoderm signal, A. *c*, In the resulting blastula the dorsal mesoderm, called the 'organizer' (O), generates a signal that dorsalizes neighbouring ventral mesoderm, C. *d*, The eventual structure of the embryo in transverse section. The signals are given different labels but they may not be different substances. They could be different amounts of the same substance, or combinations of substances.

oocyte complementary (c) DNA library, a clone homologous to bovine basic FGF. Because in the carboxy-terminal region 54 out of 61 amino acids are the same, this seems to be a genuine *Xenopus* FGF, and of course the mRNA must exist in the oocyte (although the clone isolated contained an intron, a 900-base-pair RNA can be detected in Northern blots, which is the proper size for a processed mRNA).

What of TGF- β ? Some years ago, Melton and his associates¹⁰ reported the isolation of several cDNA clones complementary to mRNAs that are localized in unfertilized eggs. One of these, Vg-1, is localized to the vegetal pole of the egg. Weeks and Melton⁶, in a paper which accompanies that of Kimelman and Kirschner, show that Vg-1 is a member of the TGF- β family.

The TGF- β family is another diverse group of cell-interaction molecules¹¹. TGF- β itself has diverse effects on cells, sometimes stimulating cell division and sometimes inhibiting it, and also has effects on the synthesis of certain specialized cellular proteins. In mammals, other members of the family include inhibin, which inhibits follicle stimulating hormone secretion, and Müllerian inhibiting

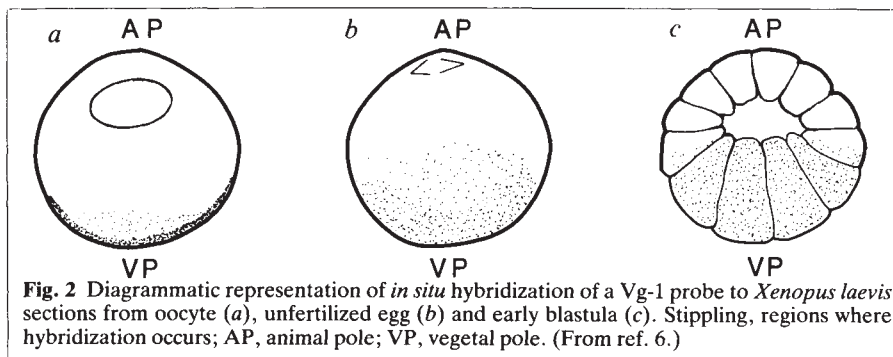


Fig. 2 Diagrammatic representation of *in situ* hybridization of a Vg-1 probe to *Xenopus laevis* sections from oocyte (a), unfertilized egg (b) and early blastula (c). Stippling, regions where hybridization occurs; AP, animal pole; VP, vegetal pole. (From ref. 6.)

embryos that have long been sought. Such determinants are presumed to give vegetal cells the property of forming gut, and animal cells that of forming epidermis. As far as mesoderm formation is concerned, it could be that vegetal cells also contain stored mRNA encoding inducing molecules (FGF, Vg-1?), and animal cells could contain stored receptors, or their mRNAs. Because animal cells become responsive to mesoderm induction at the 32/64 stage¹⁵, this could be when the receptors are deployed. It is even possible that at the equator some cells would have both receptors and inducers, and hence would induce themselves to become mesoderm. Formally, this would be equivalent to mesoderm formation by an intrinsic determinant, but by the same mechanism as mesoderm formation by cell interaction.

Future prospects

The next task is to establish whether FGF and Vg-1 mRNAs and their encoded proteins are active in the embryo. This will involve expressing them in the wrong cells, and also attempting to suppress their

normal function. The latter experiments would show if these mRNAs have an indispensable function in mesoderm formation. In this regard, it is tragic that *Xenopus* is not only unamenable to genetic analysis, but also (unlike *Drosophila*) that it contains a system which prevents antisense RNA from inhibiting translation of its complementary mRNA¹⁶. □

1. Bender, W. & Peifer, M. *Cell* **50**, 519–520 (1987).
2. Steward, R. *Science* **238**, 692–694 (1987).
3. Greenwald, I. *Cell* **43**, 583–590 (1985).
4. Wharton, K. A. *et al.* *Cell* **43**, 567–581 (1985).
5. Kimelman, D. & Kirschner, M. *Cell* **51**, 869–877 (1987).
6. Weeks, D. L. & Melton, D. A. *Cell* **51**, 861–867 (1987).
7. Vincent, J.-P. & Gerhart, J. C. *Dev Biol.* **123**, 526–539 (1987).
8. Dale, L. & Slack, J. M. W. *Development* **100**, 279–295 (1987).
9. Slack, J. M. W. *et al.* *Nature* **326**, 197–200 (1987).
10. Rebagliati, M. R. *et al.* *Cell* **42**, 769–777 (1985).
11. Massagué, J. *Cell* **49**, 437–438 (1987).
12. Jones, E. A. & Woodland, H. R. *Development* **101**, 23–32 (1987).
13. Woodland, H. R. *Biosci. Rep.* **2**, 471–491 (1982).
14. Kao, K. R. *et al.* *Nature* **322**, 371–373 (1986).
15. Jones, E. A. & Woodland, H. R. *Development* **101**, 557–563 (1987).
16. Rebagliati, M. R. & Melton, D. A. *Cell* **48**, 599–605 (1987).

Hugh Woodland and Liz Jones are in the Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK.

Mathematics

The ultimate in undecidability

Ian Stewart

"IN mathematics", said David Hilbert, "there is no *ignorabimus* . . . We must know, we will know." Like many a statement made by an ageing scientist of immaculate reputation, this one was completely wrong. Within a few years Kurt Gödel had demonstrated that within any formal system of arithmetic there are true theorems that cannot be proved, and propositions whose truth is undecidable. Indeed, it is impossible to prove that the axioms for arithmetic are not self-contradictory.

Gödel's ideas led to a new branch of mathematical logic, the theory of undecidability. Now Gregory Chaitin, of the IBM Thomas J. Watson Research Center, has proved the ultimate in undecidability theorems (*Advances in Applied Mathematics* **8**, 119–146; 1987, and *Algorithmic Information Theory*, Cambridge Univer-

sity Press, 1987), that the logical structure of arithmetic can be random. His methods are a fascinating fusion of mathematical logic and computer science. In particular, Alan Turing's analysis of the logical structure of the computational process, fundamental to both areas, plays a key role.

Hilbert emphasized the need to base the logical foundations of mathematics on a formal system of explicit assumptions, or axioms. All mathematical theorems should be logical consequences of these axioms. An analysis of the laws of deduction should lead to rigorous proofs that every true statement has a proof, and that the formal laws of arithmetic are not self-contradictory.

Gödel demolished Hilbert's programme using a clever reformulation of the hoary logical paradox 'this statement is false'. He used a coding method to model the

logical structure of an axiom system for arithmetic, within arithmetic itself. With this as an intermediary he constructed a statement in arithmetic which, interpreted via the code, effectively says 'this statement has no proof'. If the statement is true, then as it says, it has no proof. But if the statement is false, then it does have a proof — and that means it must be true.

Chaitin replaces the black-and-white dichotomy of truth versus falsehood by a scale of shades of grey, representing the computational difficulty of a proof. Instead of asking 'is this statement true?' he asks 'how complicated must a proof be?'

Information theory

The rigorous techniques for Chaitin's approach come from algorithmic information theory. An algorithm is a well-defined computer program, a list of instructions in a formally defined language. Information theory is the study of the information content of messages. One unit of information is a bit (binary digit), and a single bit provides the information to answer a single yes–no question. A message containing n bits of information can decide between 2^n equally likely possibilities. So, assuming rain or shine is equally likely, the message 'it is raining' conveys one bit of information, whereas 'it is Friday' conveys $\log_2 7 = 2.8$ bits. From Chaitin's viewpoint, Gödel's proof takes a very natural form: the theorems deducible from an axiom system cannot contain more information than the axioms themselves do.

Perhaps the most original of Chaitin's ideas is an information-theoretic characterization of randomness. Any message can be coded as a series of binary digits 0 or 1. A message consisting of n such digits carries n bits of information. But the information in some messages is compressible, in the sense that it can be encoded in a simpler form. For example, a million-digit sequence of the form 010101...01 can be described in words as 'half-a-million repetitions of 01' and this phrase can in turn be encoded (by a standard computer code such as ASCII) as a sequence of, say, 250 binary digits.

Turing's model of the computational process can be formulated entirely in terms of the manipulation of sequences of binary digits, representing input data, output data or the computer program that is manipulating the data. Indeed, if 'no electrical current' is labelled 0 and 'some electrical current' is labelled 1, this is precisely how a digital computer works. A sequence of digits can therefore be defined as non-random if it can be generated by a computer program consisting of a shorter sequence of digits, and as random if it cannot. In other words, a sequence with a pattern has a compact description, but the most efficient way to specify a random sequence is to list the sequence

itself: its information content is incompressible.

A sequence of digits 0 and 1 can be interpreted as the binary expansion of a real (non-integer) number, by inserting a 'decimal point' (or rather, a 'binary point') in front. Thus the infinite repeating sequence 010101... corresponds to the real number .010101... in binary; that is $\frac{1}{3}$. A real number is computable if its binary expansion can be the output of a computer program. It is a theorem — at first sight surprising, but true, and even easy — that 'almost all' real numbers are not computable. For example, Turing used an argument that goes back to Georg Cantor to prove the existence of at least one non-computable real number.

The proof is based on the idea that all possible computer programs, each of which must be represented by a finite sequence of digits, can be listed in order. To do this, interpret the program's defining sequence as a whole number, expressed in binary, and arrange these numbers in increasing numerical order. Now assign to each program in this list its output data, a real number expressed in binary. Run down the diagonal of this table of numbers, changing the n th digit in the n th number. The new diagonal is a number that is not on the list, which therefore corresponds to the output of no computer program whatsoever.

Turing machine

Another of Turing's basic discoveries is that it is possible to construct a universal Turing machine — a computer program capable of simulating any other program. Consider the following 'halting' problem for a universal Turing machine: given a particular input, does the computation eventually stop or does it go on for ever? Using the above 'diagonal' argument, Turing showed that the halting problem is logically undecidable.

Correspondingly, Chaitin defines a real number Ω which can be thought of as the probability that a universal Turing machine, given a random program, will eventually halt. It is impossible to compute Ω because its digits solve the halting problem for the universal Turing machine. Chaitin shows that any formal system of axioms can yield only a finite number of (scattered) digits of Ω . Thus Ω is a random real number in the above sense: it has no compact computable description.

The final step is to recast this circle of ideas in terms of whole numbers, rather than arbitrary real numbers. Chaitin does this using 'exponential diophantine' equations. These are equations, to be solved in positive integers, that can be built up from the usual algebraic operations of sums, differences and products, together with exponentiation, x^n .

Over the past two decades, logicians have solved a longstanding problem,

showing that any computation can be encoded in the set of solutions of some diophantine equation. There is a diophantine formula, for example, to generate the primes. An undecidable computation — such as the halting problem for a universal Turing machine — leads to an undecidable diophantine equation; that is, an equation in arithmetic for which there can be no proof that solutions exist, and no proof that they do not.

Finally, the randomness. Chaitin's main theorem asserts that it is possible to construct a specific exponential diophantine equation, of the form $L(n)=R(n)$, where L and R are functions that depend on a finite number of additional variables, with the following property. The equation has infinitely many whole-number solutions if and only if the n th digit of Ω is 1. Because Ω is random, the question 'does $L(n)=R(n)$ have infinitely many solutions?' has an answer which varies randomly as n takes the values 0,1,2,3,... in turn.

In the proof, Chaitin constructs an exponential diophantine equation in 17,000 variables, about 900,000 symbols long. Instead of Turing's formulation of computation, he uses a version of the standard computer language Lisp. The practical and theoretical aspects of computation become almost inseparable in this work.

For the foundations of mathematics, and even the philosophy of its application to science, this century has been one of shattered illusions. Cosy assumption after cosy assumption has exploded in mathematicians' faces. The assumption that the formal structure of arithmetic is precise and regular turns out to have been a time-bomb, and Chaitin has just pushed the detonator. His book ends with some speculations on a possible analogy with biological complexity. Resemblances to Deep Thought in Douglas Adams' *The Hitchhiker's Guide to the Galaxy* are strong, but presumably accidental. In the following quotation from Chaitin's book, the phrase 'oracle for the halting problem' boils down to 'complete specification of the number Ω ': "We have seen that Ω is about as random, patternless, unpredictable and incomprehensible as possible ... However, with computations in the limit, which is equivalent to having an oracle for the halting problem, Ω seems quite understandable ... Biological evolution is the nearest thing to an infinite computation in the limit that we will ever see: it is a computation with molecular components that has proceeded for 10^9 years in parallel over the entire surface of the earth ... Perhaps biological structures are simple and easy to understand only if one has an oracle for the halting problem." □

Ian Stewart is at the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Daedalus

Warm wind

THE temperature of the atmosphere drops with increasing altitude. Its adiabatic lapse-rate of $6.5^\circ\text{C km}^{-1}$ is maintained by its turbulence: rising air expands and cools while descending air is compressed and heated. Other gases have different adiabatic lapse-rates. The record is held by xenon, an atmosphere of which would establish a gradient of 63°C km^{-1} .

Daedalus plans to exploit this in a new wind-powered heating system. Imagine, he says, a flexible plastic bag rising a kilometre into the atmosphere, and filled with xenon. The turbulence and buffeting of the wind would soon stir the xenon into full convective equilibrium, when the bottom of the bag would be 63°C hotter than the top. If the top were maintained at local atmospheric temperature by a suitable heat-exchanger, domestically useful heat at about 70°C could be extracted from the bottom of the bag.

This ingenious scheme extracts energy, not from the 'd.c.' component of the wind, but from its turbulent 'a.c.' component. At ground-level, especially near buildings or in regions of irregular topography, much of the wind's energy is gusty and fluctuating and unavailable to conventional wind-turbines. So the thermal wind-bag has great promise.

There are two main problems. First, a flabby bag a kilometre high is rather an awkward object. Second, such a tall column of a dense gas like xenon would exert greater than atmospheric pressure at its base. This would hold the plastic bag taut, reducing its ability to flap and flex and stir the gas within. Thermal wind-bags only a hundred metres high would be more practical; they could then be anchored to buildings or lattice-towers, or even steep hillsides. If excess base-pressure is then still troublesome, Daedalus will replace his xenon by a mixture of argon and neon with the same density as the atmosphere, but an adiabatic lapse-rate about twice as great: 14°C km^{-1} . All this will reduce the temperature-rise at the bottom of the bag to only a few degrees, so Daedalus plans to multiply it up by regenerative heating. A counter-current heat-exchanger inside the bag will transfer most of the bottom heat straight up to the top again. The convective process will therefore increase the temperature at the bottom still more, and this too will be returned to the top. With a sufficiently high recycle-ratio, any desired temperature could be attained at the bottom of the bag, though with a corresponding reduction of extracted power per effective cycle. Even so, the more the cold wind buffeted your house, the more heat it would unwittingly supply! David Jones

Anomalous properties of adenine-thymine tracts

SIR—Nelson *et al.*¹ reported the single-crystal structure of DNA dodecamer containing an oligo(dA)·oligo(dT) tract. This tract is a model of (dA)_n·(dT)_n that apparently differs from ordinary B-type DNA by some unusual properties, for example: (1) conformational stability; (2) oligo(dA)·oligo(dT) abutting other sequences causes DNA bending; and (3) poly(dA)·poly(dT) does not associate into nucleosomes. What is the origin of these anomalous properties? The authors emphasize several distinctive structural features of the oligo(dA)·oligo(dT) tract: a high propeller twist (improving the intrastrand stacking); a narrow minor groove; and a system of bifurcated hydrogen bonds in the major groove, additional bond formation being promoted by the zero tilt angle. The first two features have been widely discussed², whereas the bifurcated bonds were revealed only recently^{1,3}.

It seems that experiments with duplexes, where some dA·dT pairs are replaced by dI·dC pairs, could clarify the relative role of bifurcated bonds in properties (1) and (2). Duplexes with alternating poly d(A-I)·poly d(C-T) sequences obviously cannot contain bifurcated bonds. Nevertheless, in fibres this polymer has the same B'-conformation as poly(dA)·poly(dT) and similarly is not converted into other structural forms⁴. The contribution of bifurcated bonds to DNA bending can be estimated from the data of Diekmann *et al.*⁵. When the A₅ tract is replaced by AIAIA which is accompanied by a loss of this particular system of hydrogen bonds, the anomalous electrophoretic mobility does not disappear, but decreases by about 20 per cent when A₅ is replaced by AAIAA^{3,6}, though here there is no tract of at least three successive adenines necessary to stabilize the bifurcated interaction according to Nelson *et al.*¹. The third anomalous property was explained¹ by an increased rigidity of the poly(dA)·poly(dT) caused by good stacking and bifurcated bonds. However, there is experimental evidence that poly(dA)·poly(dT) has approximately the same rigidity as poly d(A-T)·poly d(A-T) and less than poly d(G-C)·poly d(G-C) or a random DNA (see ref. 7), whereas in the latter three cases DNA can associate into nucleosomes.

Thus, the bifurcated bonds could hardly contribute much to the formation of DNA bending and conformational stability of polynucleotides in fibres. Two other factors, the overall energy balance (in particular, base-pair stacking) and the minor groove spine of hydration stabilizing the structure, suggest a qualitative explanation of several anomalies observed in (dA)_n·(dT)_n tracts^{2,8}. The relationship between the polymer rigidity and the inability

to associate into nucleosomes is not clear at present. Thus, additional experimental data are required for a final answer to the question about the determinants of (dA)_n·(dT)_n anomalous properties.

V.P. CHUPRINA

Research Computing Centre,
USSR Academy of Sciences,
Pushchino, Moscow Region, USSR

R.A. ABAGYAN

Institute of Molecular Biology,
USSR Academy of Sciences,
Moscow, USSR

1. Nelson, H.C.M., Finch, J.T., Luisi, B.F. & Klug, A. *Nature* **330**, 221–226 (1987).
2. Dickerson, R.E., Kopka, M.L. & Pjura, P. in *Biological Macromolecules and Assemblies* (eds Jurnak, F.A. & McPherson, A.) Vol. 2, 38–126 (Wiley, New York, 1984).
3. Coll, M., Frederick, C.A., Wang, A.H.-J. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8385–8389 (1987).
4. Leslie, A.G.W. *et al.* *J. molec. Biol.* **143**, 49–72 (1980).
5. Diekmann, S., Kitzing, E.V., McLaughlin, L., Ott, J. & Eckstein, F. *Proc. natn. acad. Sci. U.S.A.* (in the press).
6. Koo, H.-S. & Crothers, D.M. *Biochemistry* **26**, 3745–3748 (1987).
7. Hogan, M.E. & Austin, R.H. *Nature* **329**, 263–266 (1987).
8. Chuprina, V.P. *Nucleic Acids Res.* **15**, 293–311 (1987).

NELSON AND KLUG REPLY — Chuprina and Abagyan omit to mention the main result of our paper¹, which is that the oligo(dA)·oligo(dT) tract is straight — there is no roll nor tilt at each base pair. This is relevant to the question of how free DNA containing phased runs of adenines is bent²: the implication is that bending must occur outside the adenine tracts. We then look at the structural features which might account for the unusual properties of poly(dA)·poly(dT). We find a high degree of propeller twist between the AT base pairs, which enables a system of bifurcated hydrogen bonds to be formed and leads to the very good base-stacking interactions which occur. (Here, of

course, we cannot distinguish cause and effect.) Note that these stacking interactions would be only minimally affected by the introduction of an inosine into the adenine tracts, as an IC base pair, with only two hydrogen bonds, could accommodate a high degree of propeller twist. We have already noted that isolated examples of bifurcated hydrogen bonds have been found³.

Two other points are raised. First, the experiments on rigidity of homopolymer poly(dA)·poly(dT)⁴, quoted by Chuprina and Abagyan, rely on intercalation of dyes to determine the torsional and bending stiffness of the DNA. Not only do these dyes bind less tightly to the highly propeller-twisted AT-rich DNAs, but they also would force a decrease in the propeller twist that would destroy many of the special properties of poly(dA)·poly(dT). What is needed is a reliable determination of the persistence length of homopolymer poly(dA)·poly(dT) which would settle the question of its 'conformational rigidity'. Second, we have no evidence for the spine of hydration postulated by Chuprina to exist in the minor groove of homopolymer poly(dA)·poly(dT)⁵. This must await a higher-resolution X-ray analysis.

H.C.M. NELSON

A. KLUG

MRC Laboratory of Molecular Biology,
Cambridge CB2 2QH, UK

1. Nelson, H.C.M., Finch, J.T., Luisi, B.F. & Klug, A. *Nature* **330**, 221–226 (1987).
2. Koo, H.-S., Wu, H.-M. & Crothers, D.M. *Nature* **320**, 501–506 (1986).
3. Prive, G.G. *et al.* *Science* **238**, 498–504 (1987).
4. Hogan, M.E. & Austin, R.H. *Nature* **329**, 263–266 (1987).
5. Chuprina, V.P. *Nucleic Acids Res.* **15**, 293–311 (1987).

Open reading frames and translational control

SIR—Bürglin *et al.* have pointed out the striking sequence homologies in the 5' regions of mouse and *Xenopus* homoeobox messenger RNAs¹, especially between the Hox 1.1 RNAs from mouse² and *Xenopus* (Xhox-36, ref. 3) and the Hox 2.3 RNAs from mouse⁴, *Xenopus* (X1Hbox2, ref. 5) and man (HHO.C1, ref. 6). The importance of this sequence is indicated by the fact that it was conserved between different species as well as during a (probable) gene duplication event leading to the two closely similar genes, Hox 1.1 and Hox 2.3, in one species. Although Bürglin *et al.* have denied the significance

of open reading frames (ORFs) in this region, we would like to present evidence suggesting the importance of these ORFs and their possible role in a translational control mechanism.

During the study of murine Hox 1.1 complementary DNA sequences extending further in the 5' direction, we observed the conservation of an ORF between the murine and the *Xenopus* Hox 1.1 genes. The 5' region of the murine Hox 1.1 RNA encodes a 23-amino-acid peptide starting at position –145 (counting backwards from the ATG codon). The *Xenopus* Hox 1.1 (Xhox-36) contains an ORF at position

ATG	GGA	CAG	TCA	GAT	GGT	GGC	AGA	TCA	CGT	GGC	CCA	GGC	AGG	CAG	CTC		<i>Xenopus</i> Hox 1.1
ATG	---	---	---	---	---	TGG	CGG	TCA	CGT	GCC	GCG	GCG	AGC	---	---		<i>Mouse</i> Hox 1.1
Met	---	---	---	---	---	Trp	Arg	Ser	Arg	Ala	Ala	Ala	Ser	---	---		<i>Mouse</i> Hox 1.1
Met	Gly	Gln	Ser	Asp	Gly	Gly	Arg	Ser	Arg	Gly	Pro	Gly	Arg	Gln	Leu		<i>Xenopus</i> Hox 1.1

AGT	GTA	AAG	GAA	AAA	ATG	GGG	TTT	TGC	GTA	AAT	GTG	GGG	GTT	TAG		<i>Xenopus</i> Hox 1.1
TCC	GTC	CAA	AAG	AAA	ATG	GGG	TTT	GGT	GTA	AAT	CTG	GGG	GTG	TAA		<i>Mouse</i> Hox 1.1
Ser	Val	Gln	Lys	Lys	Met	Gly	Phe	Gly	Val	Asn	Leu	Gly	Val			<i>Mouse</i> Hox 1.1
Ser	Val	Lys	Glu	Lys	Met	Gly	Phe	Cys	Val	Asn	Val	Gly	Val			<i>Xenopus</i> Hox 1.1

Conserved open reading frames upstream of the homoeobox-containing frame in the Hox 1.1 genes of mouse² and *Xenopus* (Xhox-36, ref. 3). The distance between the stop codon and the homoeobox reading frame initiator codon is 73 and 75 base pairs, respectively.

–168 encoding a 30-amino-acid peptide³. When one considers the deletion of 21 nucleotides (7 amino acids) not only is the location of the ORF conserved, but also the structure of the encoded peptide. As shown in the figure, 14 out of 23 positions are identical, 3 conservative changes (Gly/Ala, Val/Leu) have occurred and a lysine residue is located in a central position of both peptides. On the other hand, the nucleotide sequences upstream of the ATG are totally divergent.

In the murine Hox 1.1 gene, the discussed upstream reading frame is terminated by a stop codon which overlaps with another ATG initiating a second reading frame, which again terminates with a stop codon overlapping with the ATG initiating the homoeodomain protein frame. This unusual arrangement is particularly suitable for a translational control mechanism, as translation of the first or second frame will inhibit translation of the homoeodomain-containing frame⁷. Taking into account the described evidence from the Hox 1.1 genes, and considering also the presence of upstream ORFs of different primary structure on vertebrate homoeobox RNAs such as the murine Hox 1.3 (ref. 8,9), Hox 3.1 (G. Breier, unpublished data) and Hox 1.1, the human C13 and C8 (ref. 6), and the *Xenopus* X1Hbox2 (ref. 5) and Xhox-36 (ref. 3), a translational control mechanism may well involve upstream ORFs and, at least in Hox 1.1, the encoded peptide may be important enough that it was conserved between frog and mouse.

MICHAEL KESSEL

PETER GRUSS

Department of Molecular Cell Biology,
Max Planck Institute of Biophysical
Chemistry,
Am Fassberg, 3400 Göttingen, FRG

1. Bürglin, T. R., Wright, C. V. E. & De Robertis, E. M. *Nature* **330**, 701–702 (1987).
2. Kessel, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 5306 (1987).
3. Condie, B. G. & Harland, R. M. *Development* **101**, 93–106 (1987).
4. Meijlink, F. *et al. Nucleic Acids Res.* **15**, 6773–6786 (1987).
5. Wright, C. V. E., Cho, K. W. Y., Fritz, A., Bürglin, T. & De Robertis, E. M. *EMBO J.* **6**, 4083–4094 (1987).
6. Simeone, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4914–4918 (1987).
7. Kozak, M. *Nucleic Acids Res.* **15**, 8125–8146 (1987).
8. Odenwald, W. F. *et al. Genes Dev.* **1**, 482–496 (1987).
9. Fibi, M. *et al. Development* **102**, (1988).

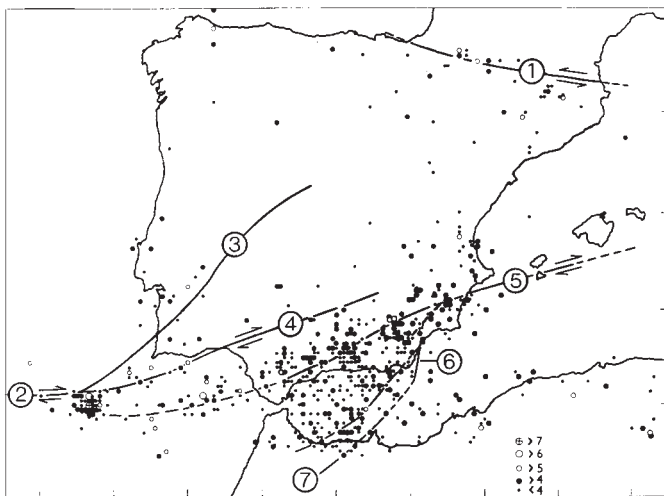
Where does Africa meet Europe?

SIR—The Straits of Gibraltar have served mankind many millenia as a morphological milestone separating different continents. To most earth scientists the Seismic Azores Transform Fault, which offsets the spreading pattern in the Atlantic Ocean by about 300 km (ref. 1), extends through the Straits of Gibraltar and separates the African and Eurasian plates in the western Mediterranean. But it is time to reappraise this view, not least because of the consequences for earthquake prediction in Spain and North Africa.

The assumption that the two plates impinge in the Straits is based on a preliminary interpretation of the plate boundary using a scattered distribution of earthquake foci². A recent higher resolution study of earthquake foci³ suggests that the Azores Fault does not continue into the Straits of Gibraltar and that no neat linear plate boundary can be identified there³. This conclusion confirms previous allegations made on the basis of the diffuse nature of seismic activity in the Straits and the adjoining Alboran Sea⁴.

Although high-resolution plots of earthquake epicentres can sometimes limit the location of major active strike-slip faults, the current deformation in the western Mediterranean is apparently distributed over so many faults that detailed field mapping provides a better means to achieve this. I have recently attempted to correlate the detailed maps and field observations made by several generations of geologists of a variety of nationalities⁵. As shown in the figure, compilation of the traces of major lineaments I have so far mapped in the western Mediterranean⁶ together with the seismic data⁴ suggests that the Azores fault passes east into a diffuse zone of splay faults in southern Spain and the Alboran Basin. The creation of the Straits of Gibraltar themselves has now been explained simply by local subsidence and erosion of the Betic-Rif Orogen or Arc of Gibraltar during the late Messinian only five million years ago⁶. This subsidence also terminated the so-called Messinian salinity crisis, which involved the deposition of up to one kilometre of salt onto the floor of the Mediterranean Sea between five and seven million years ago⁶.

A previous alternative explanation of the tectonic boundary between Europe and Africa also appears to be untenable. High-pressure, low-temperature metamorphism in the deeper nappes of the Betic Cordilleras had been used to argue that the suture between the African and Eurasian plates may be a subduction zone hidden within southern Spain⁷. But recent investigations suggest that the peak of this metamorphism in the Betic Cordillera occurred 80–85 million years ago⁸, some 60 million years before the formation of the Betic-Rif Mountains⁵. This implies that the ancient subduction zone suggested



Distribution and magnitude of earthquakes recorded by the Spanish Seismological Survey (Madrid)⁴ during the period 1961–72. Major lineaments mapped⁵ so far are: 1, North Pyrenean Fault; 2, Azores Transform Fault; 3, Plasencia lineament; 4, Guadalquivir lineament; 5, Crevillente Fault; 6, Palomares Fault; and 7, Nekor Fault.

by the exposed high-premure low-temperature rocks relates to a former plate boundary but not to the present-day suture.

I have recently suggested that the North Pyrenean Fault forms the modern southern boundary of the stable foreland of western Europe, and that the Hayes–South Atlas Fault is the northern boundary of the stable part of the African plate⁵. The area between these faults can be interpreted as the Iberian microplate which indents Italy in a fashion similar to the impingement of the Indian plate with Tibet before continental collision. This proposition would reduce the size of the African plate by 2 million km², but needs careful testing in the future.

Unfortunately the only research group studying active faults in the western Mediterranean does not include the South Atlas Fault in its programme⁹. Details of the geology, seismic data and aeromagnetic anomalies are still unclear. Such studies are important not only for solving the complex nature of the plate boundary but also to increase the accuracy of predicting earthquakes that might put Mediterranean people, particularly in Spain, Morocco, Algeria and Tunisia, at risk.

RUUD WEIJERMARS

The Hans Ramberg Tectonic Laboratory,
Institute of Geology, Uppsala University,
Box 555, S-751 22 Uppsala, Sweden

1. Rona, P. A. NOAA Atlas 3 (US Dept. of Commerce, National Oceanic and Atmospheric Administration, Miami, 1980).
2. McKenzie, D. P. *Nature* **226**, 239–243 (1970).
3. Grimson, N. L. & Chen, W. P. *J. geophys. Res.* **91**, 2029–2047 (1986).
4. Udias, A., Arroyo, A. L. & Mezcu, J. *Tectonophysics* **31**, 259–289 (1976).
5. Weijermars, R. *Geol. Rundschau* **76**, 667–676 (1987).
6. Weijermars, R. *Tectonophysics* (in the press).
7. Platt, J. P. & Behrmann, J. H. *J. struct. Geol.* **8**, 15–33 (1986).
8. De Jong, K. *Conf. Geol. kinemat. dynamics (Abstr.) Uppsala*, 22–24 April 1987.
9. Hancock, P., Boccaletti, M. & Moratti, G. *J. struct. Geol.* **9**, 255–257 (1987).

Molecular recognition in biomineralization

Stephen Mann

School of Chemistry, University of Bath, Bath BA2 7AY, UK

The deposition of precise arrays of inorganic crystals in many organisms involves controlled nucleation at interfaces between the crystals and substrate macromolecules. These inorganic-organic molecular recognition processes have potential technological application.

SURPRISING though it may seem inorganic solid state chemistry is as much an integral aspect of biology as the aqueous phase organic and inorganic reactions of conventional biochemistry. More than 40 different minerals have been identified in organisms (Table 1), the most common being the phosphate and carbonate salts of calcium that are used in conjunction with organic polymers such as collagen and chitin to give structural support to bones and shells. Salts of barium, strontium, silicon and iron are also common. The minerals are formed by normal cellular processes and replicate to produce precisely organized structures in cells ranging from bacteria, algae and protozoa to the osteoblasts of bone. The minerals may be in membrane-bound vesicles within cells, in the mucilaginous layers of cell walls in bacteria, or impregnated in biopolymers in extracellular space. In all these structures, the inorganic crystals are laid down in orderly arrays in association with a matrix of organic macromolecules. There is a growing body of information on the structure, composition and synthesis of the macromolecules, but the key and exciting question is to identify the molecular processes that produce minerals of precise form with uniform particle size, novel crystal morphology and specific crystallographic orientation. How do the proteins and polysaccharides of the substrate interact to regulate solid-state chemical reactions?

Distribution

Calcium salts are not the only minerals found in biological systems, and minerals serve several functions other than providing support. Many unicellular organisms build cytoskeletons by depositing hydrated silica, a form of amorphous glass; in protozoa and algae this inorganic polymer is often sculpted into exquisite rods, perforated disks and reticular frameworks that are species-specific and have long been used as diagnostic characters for classification. Minerals are also put to special uses, for example, as magnetic sensors in magnetotactic bacteria (Fe_3O_4), gravity balance devices (CaCO_3 , CaSO_4 , BaSO_4), deterrence against predators (SiO_2 , CaCO_3), iron storage and mobilization ($\text{Fe}_2\text{O}_3 \cdot n\text{H}_2\text{O}$ in the protein ferritin), love darts in snails (CaCO_3) and eye lenses (CaCO_3 in fossilized trilobites). The hard parts formed by these minerals are living structures that may undergo active demineralization and remodelling in response to environmental and biological stress. Bone, for example, provides an essential store of calcium in vertebrates and calcium oxalate is a calcium source in plants.

Controlled crystallization

Because the interactions resulting in controlled crystallization are tailored to biological function, the underlying molecular processes must have a specificity analogous to those involved in the reactions of aqueous-phase biological molecules. An understanding of biological solid-state interactions would therefore be of immense value in structural biology and medicine (for example, in the pathological mineralization of bones and teeth and formation of kidney stones), in crystal growth, colloidal and solid-state science (as in the prevention of industrial scaling and the controlled synthesis of electronic, magnetic and

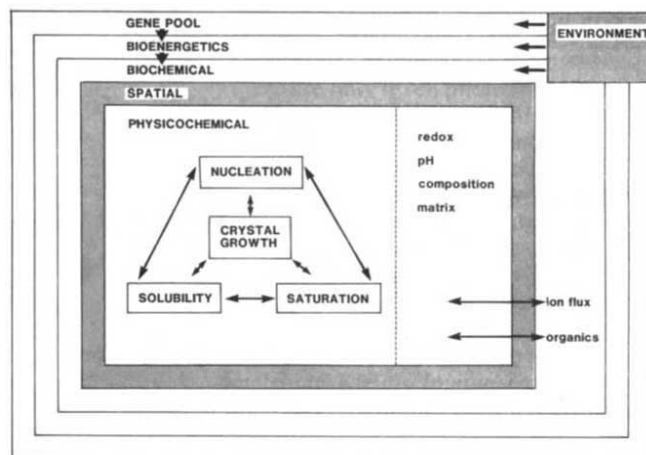


Fig. 1 The control processes in biomineralization³. In organisms, there are several different interconnecting levels that regulate the physical chemical properties of mineralization (solubility, supersaturation, nucleation and crystal growth). An essential condition for controlled mineralization is spatial localization arising from the compartmentalization of biological space. This permits direct regulation of physicochemical and biochemical properties in the mineralization zone. Nucleation, in particular, can be mediated by organic polymeric substrates in or on the spatial boundary. At a higher level of organization, mineralization is under biochemical and bioenergetic constraints and, ultimately, under control at the gene level. The interplay of these control processes with the external environment is also of fundamental importance.

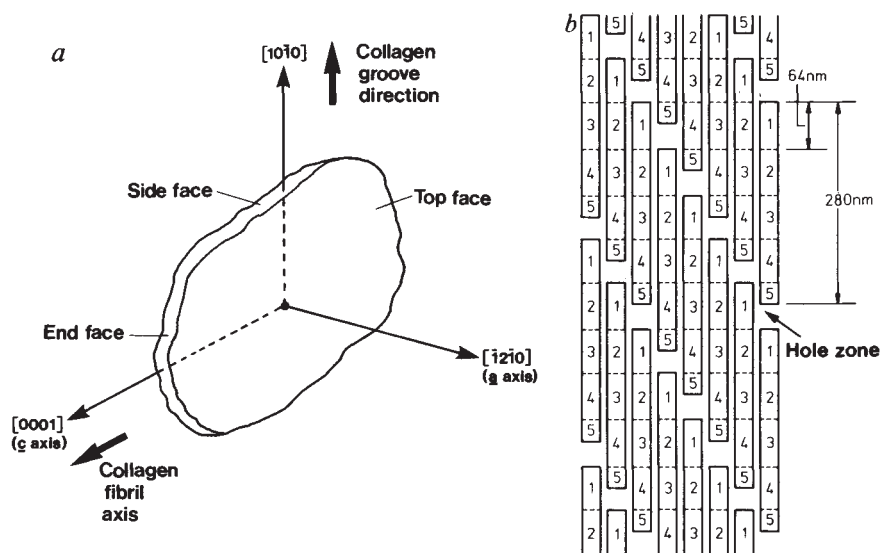
catalytic devices) and in materials and engineering technology (organized composites and ceramic precursors, and the interrelationships between microstructure and mechanical properties, for example).

The influence of organic macromolecules is important in the regulation of growth of the mineral^{1-4,5,6} and in the resulting specificity in crystal morphology⁷ and particle aggregation⁸, but the molecular interactions inherent in these processes are not well understood. Here we will focus primarily on current ideas about the role of organic macromolecules in initiating the formation (nucleation) of inorganic solids in biology. Because these ideas are potentially important in industrial processes where nucleation control is a critical factor, for instance in the reproducible synthesis of materials, emphasis will be placed on the molecular interactions at the interface between the mineral and organic matrix. It is in the area of biological mechanisms of nucleation control that significant novel concepts have arisen, and the investigation of these ideas in simplified physical chemical systems should provide information about the regulation of nucleation in biomineralization.

Nucleation control

The processes controlling biomineralization are summarized in Fig. 1. Organized biopolymers at the sites of mineralization are

Fig. 2 The relationship between the crystallographic properties of calcium phosphate crystals of bone and the preorganization of collagen fibrils, as determined by high resolution transmission electron microscopy⁹. Numbers in square brackets refer to the Miller index notation for describing crystallographic directions. *a*, A single plate-like crystal from calcified turkey tendon showing its orientation with respect to the collagen fibrils. The crystallographic *a* $[1\bar{2}10]$ and *c* $[0001]$ axes are aligned perpendicular and parallel to the collagen fibril axis respectively. The $[10\bar{1}0]$ crystallographic direction runs parallel to the collagen groove direction. *b*, The revised quarter-stagger model for collagen (in planar form), showing the intrafibrillar regiospecific hole zones which are the sites of hydroxyapatite nucleation in mature bone. In this arrangement each molecule is divided into five zones, the first four of equal length (64 nm) and the fifth only 25 nm. Adjacent molecules are transposed by 64 nm along their length by strong intermolecular cross linking to produce a $30\text{--}40\text{ nm} \times 3\text{--}5\text{ nm}$ space (the hole zone) between the ends of each molecule. Several models of the collagen fibril arrangement predict that grooves are formed by the overlap of adjacent hole zones and that the grooves are organised in parallel rows at the same height in the fibrils^{14,15}. (Reproduced with permission from ref. 40).



essential to these processes. In unicellular organisms these macromolecules act primarily as spatial boundaries through which ions are selectively transported to produce localized supersaturation within discrete cellular compartments. In many instances, particularly in organisms such as the diatoms that deposit shells of amorphous silica, the final shape of the mineral appears to be dictated by the ultrastructure of the membrane-bound compartment. Thus a diversity of mineral shapes can be biologically moulded by constraining the space available to the growing mineral and there is no immediate reason to infer that specialized interactions between the mineral and organic macromolecules exist.

There is clear evidence in many multicellular organisms for a much closer relationship between macromolecules and crystallographic properties. Two well-studied examples are bone and mollusc shells. Bone is made up of microscopic plate-like calcium phosphate crystals ($\sim 450\text{ Å} \times 200\text{ Å} \times 30\text{ Å}$) formed within and between fibrils of collagen; in calcified turkey tendon, each crystal is deposited with the *a* axis of the crystal perpendicular to the plane of the plate, the length of each plate corresponding to the *c* axis of the crystal, which is also oriented parallel to the long fibre axis of the protein⁹ (Fig. 2*a*). There is similar relationship in the nacreous inner layers of some mollusc shells, where alternating layers of polygonal blocks of aragonite are sandwiched between thin (30–300 nm) sheets of a protein-polysaccharide organic matrix; each crystal is aligned with the *c* axis perpendicular to the plane of the organic sheets (Fig. 3*a* and *b*). Although the biopolymers in bone and shells certainly have important functional roles in the mechanical strength and toughness of the mineralized composite, the well-defined crystallographic relationships imply there is an underlying molecular function for the macromolecules. In particular, the formation of crystals with preferred orientations at specific sites on the organic matrix indicates that interactions at the mineral-matrix interface are the key to the regulation of nucleation.

Nucleation, in general, represents an activation energy barrier to the spontaneous formation of a solid phase from a supersaturated solution. This kinetic constraint may be sufficient to offset the thermodynamic driving force for precipitation, resulting in metastable solutions which do not undergo phase transformations over a long period of time. Clusters of the solid phase will continue to grow only if the energy required to form the new interface, ΔG_i , is overcome by the energy released in the forma-

tion of bonds in the bulk of the aggregate, ΔG_B . Nucleation is favoured only at a critical cluster size, which is directly proportional to the ratio $\Delta G_i/\Delta G_B$. The activation energy for nucleation (ΔG_N) is related to ΔG_i and ΔG_B by the equation

$$\Delta G_N = \frac{16\pi(\Delta G_i)^3}{3(\Delta G_B)^2}$$

and

$$\Delta G_B = kT \log_e S$$

where *k* is the Boltzmann constant, *T* the temperature and *S* the relative supersaturation of the medium.

In biological environments, the activation energy for nucleation can be reduced by lowering the interfacial energy (ΔG_i) and/or increasing the supersaturation^{1–3}. Supersaturation levels are regulated by confining the mineralization reactions to diffusion-limited sites (Fig. 1), such as intracellular vesicles and membrane-bound extracellular compartments, where selective ion-transport to and from the mineralization zone can effectively be controlled^{4,10}. Interfacial energies, on the other hand, can be lowered by the presence of organic surfaces at the nucleation site. A factor of fundamental importance is the existence of a discrete energy minimum for nucleation on the organic substrate, corresponding to a discrete nucleus structure and orientation. If there were a range of minima corresponding to the formation of nuclei of varying structure and morphology on the organic substrate, then crystallochemical specificity in biomineralization simply would not arise.

Because the chemical bonding at the surface of mineral nuclei is primarily ionic, the organic interface must contain areas of high local charge where electrostatic, dipolar and hydrogen bonding interactions can take place during nucleation. If these forces provide molecular recognition between the stereochemical requirements of ions in the nucleus surface and charged groups at the macromolecular interface, then both the structure and orientation of the mineral deposit can be precisely determined. Two features of the organic matrix are essential to generate this crystallochemical specificity: an existing organization of the organic matrix (molecular preorganization) and molecular complementarity between the inorganic ions and the local binding sites on the matrix. To limit the number of discrete nucleation sites at the matrix surface, the matrix must be organized at both the microscopic and molecular level before mineral

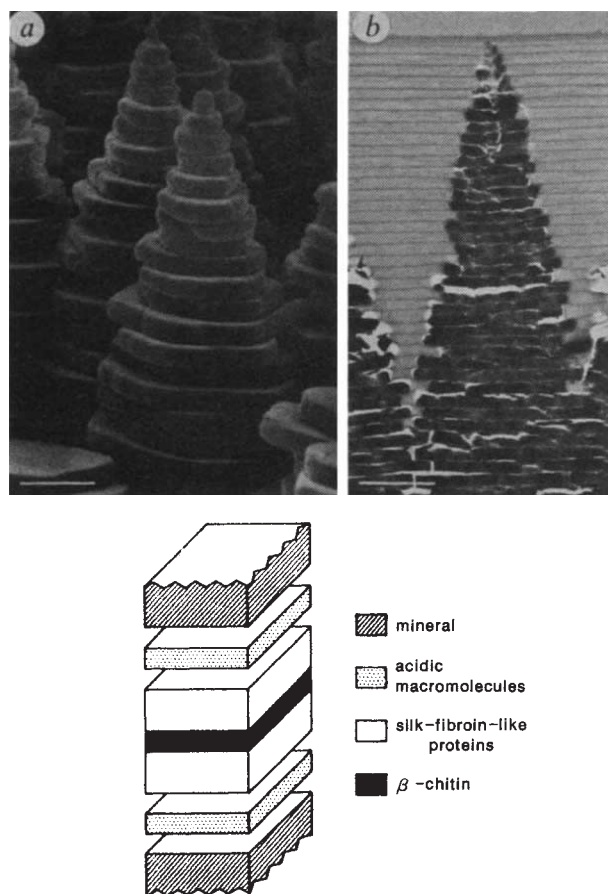


Fig. 3 The relationship between aragonite crystals and planar sheets of organic matrix in the mollusc shell. *a*, Scanning electron micrograph of the growth surface of the nacreous layer from the mollusc *Monodonta labio* showing the stacked arrangement of aragonite crystals. Scale bar, 2.5 μm . *b*, Transmission electron micrograph of an unstained section showing growing stack (*T*) with sheets of organic material (*S*) periodically interspaced between the aragonite crystals. Scale bar, 4 μm . (Reproduced with permission from ref. 41). *c*, Schematic representation of a composite section of a mollusc shell showing one sheet of organic matrix bounded by the mineral aragonite. The diagram illustrates the simplified two-component model for the organization of the matrix, with hydrophobic biopolymers (chitin and silk-fibroin proteins) forming a framework on which the hydrophilic nucleator acidic macromolecules are arranged. (Reproduced with permission from ref. 23).

deposition starts. The disposition of the nucleation centres is regiospecific in that they are spatially confined to localized domains in the polymeric structure of the organic framework; the distribution of the sites may be random or periodic (Fig. 4). The specificity of the sites for nucleation arises from the organization of the matrix to produce distinct structural, topographical and chemical domains across its surface. The interaction between ions in the crystal nuclei and the localized surface binding sites on the matrix is cooperative and results in crystal-chemical specificity in the mineral deposit. Complementarity at the inorganic-organic interface is both electronic and steric; thus the matrix nucleation centres can be considered analogous to receptor sites in enzymes, genes and antibodies, and the nucleation ion-clusters analogous to enzyme substrates, cofactors and antigens.

Matrix preorganization

The preorganization of extracellular macromolecular substrates for regiospecific nucleation and the subsequent development of

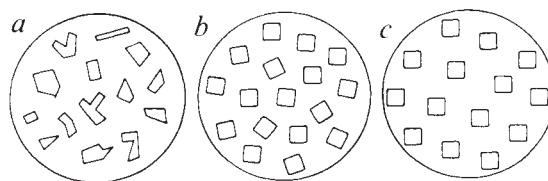


Fig. 4 Nucleation of biominerals on organic polymeric substrates. The organic surface within the area delineated by the circles consists of a limited number of discrete sites at which crystal nucleation takes place. *a*, Non-specific nucleation in which crystal nuclei of varying orientation are randomly formed across the organic surface. *b*, *c*, Site-directed nucleation involving molecular specificity at the organic/crystal interface. The crystals have a preferred orientation and morphology and may be organized randomly as shown in *b* or periodically as in *c*.

biominerals with controlled micro-architecture can be clearly demonstrated in many systems. In avian eggshells, the nucleation sites are spherulitic protein masses organized on the fibrous disulphide-linked proteins of the inner shell membranes. Oriented calcite nucleation occurs at these loci with the crystals developing preferentially along the crystallographic *c* axis. This process occurs only once in the development of the avian eggshell and is therefore confined to the two-dimensional surface of the outer membrane; in contrast, nucleation sites in mollusc shells form repeatedly throughout the lifetime of the animal through the episodic deposition of the matrix surface, and so result in three-dimensional stack-of-coins (Fig. 3*a* and *b*) or brick-wall arrangements of aragonite crystals, each with the crystal *c* axis oriented perpendicular to the matrix sheets.

The organization of nucleation at discrete sites along the organic matrix is determined by the surface and bulk structure of the substrate. These properties are dictated by the underlying molecular processes involved in the polymerization and self-assembly of the organic macromolecules. The best-studied example is collagen in bone, which comprises a rope-like arrangement of three, non-coaxial, helical polypeptides rich in glycine, proline, hydroxyproline and alanine; the small glycine side-chains at every third amino acid allow close-packing of the polypeptides. This configuration is the basic subunit for the assembly of the collagen fibre, which can be generated from a variety of packing permutations, depending on the chemical composition of the polymerization medium¹¹. This gives the collagens involved in biomineralization the specific revised quarter-stagger configuration^{12,13}, in which holes and grooves^{14,15} are organized in parallel rows across the fibrils (Fig. 2*b*). These spaces are the sites for calcium phosphate mineralization in mature bone and the localized structure, chemistry and topography of these regions must be important factors in controlling nucleation; for example, native-type reconstituted collagen fibres and collagen fibres obtained after decalcification of bone tissue can be mineralized *in vitro* with crystal formation specifically in the hole zones, whereas reconstituted collagen fibres with different (non-native) packing arrangements fail to nucleate calcium phosphate crystals under identical physical chemical conditions¹⁶.

There is, however, much discussion about the precise function of structural proteins such as collagen in biomineralization. In particular, old ideas that the molecular nature of the hole zone directly promoted calcium phosphate nucleation have to some extent been replaced by schemes that consider collagen as providing the necessary architecture (in the form of holes and grooves) for the location of non-collagenous nucleator molecules, such as glyco- and phosphoproteins. This two-component model of the organic matrix, with an essentially hydrophobic constituent that acts as an architectural framework on to which hydrophilic constituents directly involved in the control of mineral nucleation are attached (Fig. 3*c*), has been

Table 1 The types and functions of the main inorganic solids found in biological systems

Mineral	Formula	Organism/function
Calcium carbonate:		
Calcite	CaCO_3^*	Algae/exoskeletons Trilobites/eye lens
Aragonite	CaCO_3	Fish/gravity device Molluscs/exoskeleton
Vaterite	CaCO_3	Ascidians/spicules
Amorphous	$\text{CaCO}_3 \cdot n\text{H}_2\text{O}$	Plants/Ca store
Ca phosphate:		
Hydroxyapatite	$\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$	Vertebrates/endoskeletons teeth, Ca store
Octa-calcium phosphate	$\text{Ca}_8\text{H}_2(\text{PO}_4)_6$	Vertebrates/precursor phase in bone?
Amorphous	?	Mussels/Ca store Vertebrates/precursor phases in bone?
Calcium oxalate:		
Whewellite	$\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$	Plants/Ca store
Weddellite	$\text{CaC}_2\text{O}_4 \cdot 2\text{H}_2\text{O}$	Plants/Ca store
Group IIA metal sulphates:		
Gypsum	CaSO_4	Jellyfish larvae/gravity device
Barite	BaSO_4	Algae/gravity device
Celestite	SrSO_4	Acantharia/cellular support
Silicon dioxide:		
Silica	$\text{SiO}_2 \cdot n\text{H}_2\text{O}$	Algae/exoskeletons
Iron oxides:		
Magnetite	Fe_3O_4	Bacteria/magnetotaxis Chitons/teeth
Goethite	$\alpha\text{-FeOOH}$	Limpets/teeth
Lepidocrocite	$\gamma\text{-FeOOH}$	Chitons (Mollusca) teeth
Ferrihydrite	$5\text{Fe}_2\text{O}_3 \cdot 9\text{H}_2\text{O}$	Animals and plants/ Fe storage proteins

* A range of magnesium-substituted calcites are also formed.

applied to a range of biomineralized structures in both vertebrates and invertebrates. Much of the present information comes from investigations on mollusc shells. Biochemical studies of the matrix composition show that there is a complex assemblage of macromolecular constituents at the mineralization site and that these constituents may differ between different species¹⁷ and between different layers in the same shell¹⁸. But comparative studies of shells decalcified at neutral pH in the presence of EDTA indicate that the EDTA-soluble extract is remarkably similar in all species studied. It is a heterogeneous mixture of two classes of proteins¹⁹: one rich in aspartic acid and, to a lesser extent, glutamic acid (representative amino acid composition Asp 30%, Glu 17%, Ser 10%, Gly 7%, Thr 7%), possibly with some covalently bound polysaccharides, and the other rich in serine and associated with relatively large amounts of polysaccharide, possibly proteoglycans. In contrast, the EDTA-insoluble fraction from mollusc shells varies considerably in composition between species. The major components are proteins rich in glycine, alanine, phenylalanine and tyrosine²⁰.

On the basis of these results, combined with ultrastructural data, the mollusc-shell matrix is now considered to be highly organized before the nucleation of aragonite crystals (Fig. 3c). The EDTA-insoluble constituents provide an architectural framework based on a core of β -chitin sandwiched between two layers of silk-fibroin-like proteins; these are covered on both sides by layers of EDTA-soluble matrix constituents which are aligned with the underlying anti-parallel β -pleated sheet polypeptide chains^{21,22} to provide a structurally well-defined charged interface for nucleation²³. This organization of acidic and hydro-

phobic macromolecules explains both the mechanical and crystallochemical properties observed in shells.

Molecular complementarity

Having established that organic biopolymers are organized so that their surfaces consist of localized sites tailored to the control of mineral nucleation, what type of interactions are possible at the inorganic-organic interfaces?

The key to specificity in nucleation is the presence of some form of molecular complementarity at an interface between functional groups on the organic macromolecules and ions in the surface of a crystal nucleus. It is postulated that polarity, charge and stereochemical relationships at sites that are restricted in size and topology can induce matching between the potential fields surrounding the inorganic and organic surfaces. This would account for the specific lowering of the activation energy for nucleation that is observed as a function of the structure and orientation of the nuclei. In this perspective, the molecular specificity of organic surfaces in nucleation resembles that inherent in protein-protein or protein-membrane associations. Precise chemical, configurational and topographical organization of hydrophilic and hydrophobic domains at the surface of the macromolecules in each system generates similar specific changes in the conformation of multi-component complexes.

Three aspects of recognition at inorganic-organic interfaces leading to specificity in the nucleation of biominerals are of particular interest here: electrostatic accumulation, structural correspondence and stereochemical requirements. Although these factors probably act cooperatively, we will discuss them separately to highlight the essential properties of the matrix surface.

Electrostatic accumulation. Organic macromolecules extracted from shell and dentine bind Ca^{2+} *in vitro* with stoichiometries often exceeding the number of potential anionic binding sites on the matrix²⁴⁻²⁶. This is taken to indicate that the electrostatic accumulation of ionic charge at specific sites on an organic substrate—the so-called ionotropic effect²⁷—induces crystal nucleation. It is unlikely that this supra-stoichiometric binding involves multiple coordination complexes between organic ligands and Ca^{2+} because the stereochemical requirements of multiple ligands associated with distinct amino-acid sequences and molecular folding would suggest higher affinities and lower binding capacities than those reported. A more feasible interpretation is that the accumulation of charge at surface (anionic) sites on the matrix occurs by primary cation-binding. This would generate a loosely associated anion-coordination sphere and in turn attract a secondary cation-coordination sphere. This effect has been demonstrated by ³¹P NMR spectroscopy of the binding of Ca^{2+} to phosphorylated serine residues in the presence of inorganic phosphate in a phosphoprotein isolated from fetal dentine²⁸. (Phosphoproteins have also been isolated from chicken and bovine bone^{29,30}.)

Both high capacity and low binding affinity are critical for controlling nucleation. A matrix with high affinity for Ca^{2+} would probably inhibit nucleation because strong polymer binding would effectively remove aqueous ions from solution and render them unreactive. On the other hand, high capacity is required to localize a sufficient number of ions at the matrix surface to exceed the critical nucleus size. Crystallochemical selectivity will arise at the interface if the polarity and charge density at the matrix nucleation sites complement the electrostatic field generated around a particular crystal face of the forming nucleus. Important factors are the size of the nucleation site, the number of active functional groups on the matrix and their disposition, determined by the macromolecular structure and topography of the sites. For example, one model for hydroxyapatite nucleation in bone proposes that phosphoprotein molecules present in the regiospecific hole zones of collagen electrostatically accumulate large amounts of Ca^{2+} and

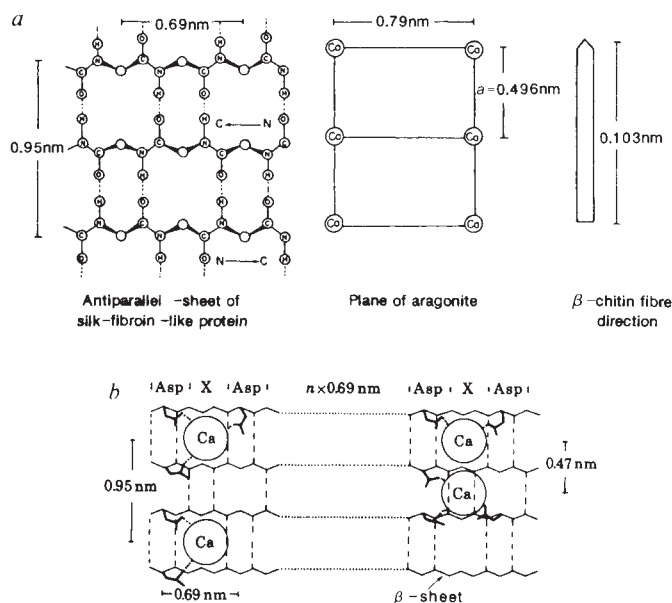


Fig. 5 Molecular correspondence at the inorganic-organic interface in the nacreous shell layer of *Nautilus repertus*. *a*, Schematic representation of the structural relationships between protein sheets, aragonite crystals and chitin fibres. There is a close geometric match between the periodicity of the β -sheet and the lattice spacings in the ab face of aragonite. *b*, Possible modes of molecular complementarity between Ca atoms in the aragonite ab face and aspartic acid residues organized in the sequence Asp-X-Asp (X, neutral residue) along the β -sheet matrix interface. Ca binding involves two or three ligands and is regulated along the interface to produce oriented crystal nuclei. (Reproduced with permission from refs 21 and 23 respectively).

phosphate ions³¹. This effect could be enhanced by the distinct size and topology of the hole zones to increase the structural order of the transient nucleus in its second and higher coordination spheres. This would facilitate deposition of hydroxyapatite rather than of a less-ordered material, such as amorphous calcium phosphate or microcrystalline brushite. Furthermore, it is feasible that the topographical organization of ion charge in the hole zone favours orientational anisotropy in the development of the mineral nuclei, resulting in the hydroxyapatite crystals being preferentially aligned, with the crystallographic c axis running parallel to the length of the hole zone, namely the collagen fibre axis (Fig. 2a).

Structural correspondence. Structural correspondence, or epitaxy, between the macromolecular surface and crystal nuclei implies a one-to-one geometrical matching between the functional groups of the matrix and the crystallographic lattice dimensions of a specific crystal face in the overlying mineral phase. Although structural relationships at the mineral-matrix interface have been implicated in a wide range of biomineralization processes, they have received only limited support from experimental studies. Some of the most elegant work has involved electron diffraction by small ($6 \mu\text{m}^2$) regions of partially demineralized mollusc shell fragments which in many species show that both the a and b axes of the antiparallel β -pleated sheet of the matrix and the overlying aragonite crystal lattice are matched in orientation at the interface²² (Fig. 5a). Although this crystallographic correspondence is present throughout each matrix sheet the alignment of the ab plane of the β -pleated sheets can change, with the result that the mineral layer consists of a mosaic of aragonite crystals all with their c axis perpendicular to the matrix surface, but with different crystallographical alignment of the a and b axes parallel to the matrix interface²³.

A plausible mechanism has been suggested for the crystallochemical correspondence observed in mollusc shells²¹⁻²³: the presence of a local sequence of negatively charged residues such as aspartic acid, linked to the β -sheet conformation, could act as a structurally organized nucleation site by binding Ca^{2+} in a configuration corresponding to the ab plane of aragonite. Comparison of the Ca-Ca distance in the ab plane of aragonite with the matrix periodicity indicates that a close matching occurs along the a axes ($4.96 \text{ \AA}$ and $4.7 \text{ \AA}$ respectively), whereas a greater mismatch is observed along the b axes ($7.97 \text{ \AA}$ and $6.9 \text{ \AA}$ respectively). In contrast, the periodicities along the b axes are essentially commensurate over a distance of seven calcium atoms ($48 \text{ \AA}$). As Ca^{2+} binding to carboxylate groups is generally cooperative, involving at least two or three ligands³², an amino-acid sequence of Asp-X-Asp (where X is a neutral residue) along the β -sheet framework could present optimum binding (and hence nucleation) configuration for Ca^{2+} at the matrix interface²³ (Fig. 5b). Sequences of this type are common in EDTA-soluble proteins of both aragonite and calcite layers of mollusc shells^{33,34}. Domains along the β -sheet that do not contain this sequence would be inactive in nucleation. Thus the coding of such sequences at intervals along the β -sheet framework would provide regiospecific sites for mineralization over considerable distances in extracellular space.

Stereochemical requirements. Structural correlations (as described above) depend essentially on numerical relationships in one or two dimensions at the mineral-matrix interface. Although some consideration of cation coordination geometry is implied by specific amino-acid sequences, these factors are secondary because of the flexible stereochemistry of Ca-ligand interactions. No emphasis is placed, however, on the stereochemical requirement for optimal completion of the anionic coordination geometry in the developing crystal face. This is an important aspect of recognition at inorganic-organic interfaces leading to specificity in nucleation.

Many biologically deposited calcite and aragonite crystals have their crystallographic c axes oriented along preferred directions. The disposition of Ca atoms in lattice sites running perpendicular to the c axis is almost identical in both structures and an argument based on structural correspondence can be invoked for the preferential adoption of the c axis orientation in these biominerals. This mechanism fails to explain, however, why selectivity between these two crystal structures is such a common feature of biomineralized systems; for example, many shells have inner pearly layers of aragonite and outer prismatic layers of calcite. It turns out that the difference between the calcite and aragonite structures along the c axis lies in the stereochemistry of the CO_3^{2-} anions sited between adjacent Ca layers running perpendicular to the c axis. In calcite, the most favoured stereochemical approach for the three binding oxygens of carbonate is parallel to the calcium plane and the optimum face for this presentation is the ab or (001) face. It follows that a matrix surface with binding sites in a configuration that is stereochemically equivalent to the anion motif of the crystal structure provides the optimum local coordination geometry to induce oriented nucleation for a specific crystal face in the nucleus. This stereochemical requirement explains the selectivity of interfacial association between crystal faces with similar Ca-Ca distance (structurally equivalent) but with different anion stereochemistry (for example, calcite and aragonite). It also explains why faces of maximum charge density do not always interact with the matrix surface, as might be predicted for a purely electrostatic recognition at the interface.

The importance of stereochemical complementarity has been demonstrated experimentally *in vitro*⁷. Acidic protein macromolecules in the β -sheet conformation extracted from mollusc shells and adsorbed at low concentration ($<0.5 \mu\text{g ml}^{-1}$) on to rigid glass or plastic substrates developed rhombohedral calcite crystals, which were often nucleated on a small triangular (001) face at the rigid substrate surface. Indirect immunofluorescence

staining of the crystals clearly showed the presence of aspartic acid-rich protein on the (001) face only, indicating that the organization of carboxylate residues attached to the β -pleated sheet surface satisfies the important stereochemical requirements of this face.

Cooperativity is also important in the regulation of crystal nucleation on ordered organic substrates³⁵. In outline, the mechanism proposed for oriented (001) calcite nucleation involves a cooperative effect arising from enhanced electrostatic interaction between sulphonated residues of the substrate and the structural and stereochemical requirements of polyaspartate macromolecules. Evidence for this mechanism comes from recent *in vitro* experiments using polyaspartate in the β -sheet conformation adsorbed onto polystyrene films. It did not induce nucleation of oriented calcite crystals unless the substrate had previously been sulphonated. The effect is clearly cooperative because only a few oriented crystals form on sulphonated films without polyaspartate, whereas many form on films with both components, and the affinity of Ca for sulphonated films with adsorbed polyaspartate is more than twice that for sulphonated films without the polypeptide. These results are consistent with observations on acidic macromolecules isolated from mollusc shells, which contain covalently bound sulphate residues. When adsorbed on to solid substrates the macromolecules induce oriented calcite nucleation. As many mineralized tissues contain macromolecules with sulphate and carboxylate ligands, the demonstration of cooperativity in these model systems has wide-ranging implications in the understanding of nucleation in living systems.

Potential applications

The study of biomineralization is relevant to a range of industrial processes, including the direct biotechnological exploitation of biogenic materials³⁶, the design of novel composite materials by *in situ* precipitation³⁷, and the mimicking of biological processes in an inorganic context. The underlying mechanisms of controlled nucleation in biological systems has important implications for the third of these approaches, for fabricating inorganic materials with technological applications. In principle, the aim is to use stereospecific polymer surfaces, prepared either synthetically or, in the longer term, by molecular biology techniques, on which crystals of defined shape, size, structure and orientation could develop. Crystal properties would be regulated by the molecular interactions at the interface, with systematic changes in the substrate surface reflected in changes in crystal chemistry. If this could be achieved, a range of useful materials could be fabricated, including magnetic (Fe_3O_4 , $\gamma\text{-Fe}_2\text{O}_3$), catalytic (doped metal oxides), optical (SiO_2 , LiNbO_3), and piezoelectric (BaTiO_3 analogues) devices. Furthermore, the growth of materials as oriented thin sheets on polymeric substrates has potential value for making piezoelectric and elec-

troactive sensors, as well as in the production of organized precursors for chemical and ceramic support beds.

Apart from the *in vitro* work described above on the nucleation properties of β -sheet proteins attached to sulphonated polymer films³⁵, so far only a few studies have investigated the nucleation potential of organic substrates. Perhaps the most promising is the system based on compressed surfactant or phospholipid monolayers where control of both the polarity and the head-group spacing of these molecules could provide a way to engineer the interactions involved in nucleation at an organic-inorganic interface. Landau *et al.*^{38,39} have shown that oriented growth of glycine and sodium chloride takes place at the air/water interface of structured monolayers. The controlled formation of chiral crystals such as glycine is valuable because homochiral monolayers will induce the nucleation only of crystal faces with corresponding symmetry. Although the nucleation of NaCl on compressed monolayers is less specific, probably because the surfactant molecules cannot simulate the stereochemistry of the NaCl surfaces, particular faces are stabilized in the presence of charged monolayers, presumably by cooperative electrostatic and structural correspondence³⁹. For example, close-packed films of octadecylamine favour nucleation of NaCl on the cubic (100) face, whereas stearic acid monolayers induce nucleation mainly on the uni-charged (111) faces. Using stearic acid, complete inhibition of crystal nucleation under the monolayer results from increasing the distance between the carboxylate groups of the monolayer, which can be achieved by using molecules with bulky hydrophobic chains, such as cholesteryl succinate. Although only a limited number of crystal types have been studied, this is a novel approach in engineering crystal nucleation that can be extended into fields such as crystallization science and materials technology.

The significance of studying inorganic solid-state chemistry in biology is highlighted by the importance of molecular recognition at interfaces comprising organic polymeric substrates and inorganic crystal surfaces, particularly for controlling the crystallochemical properties of mineral nuclei forming at the matrix surface. The key features of the interface are that it is chemically, structurally and topographically organized for site-directed nucleation, and that the molecular interactions at each nucleation site involve a high degree of complementarity between the electrostatic, structural and stereochemical requirements of functional groups at the matrix surface and ions in the crystal faces of the developing nuclei. The *in vitro* investigation of these factors and their theoretical simulation by molecular modelling together with the investigation of possible industrial and biotechnological applications, are important directions for future work.

I thank Professor R. J. P. Williams and Dr S. Weiner for reading the draft manuscript and for valuable comments.

- Mann, S. *Struct. Bond. (Berlin)* **54**, 125-174 (1983).
- Williams, R. J. P. *Phil. Trans. R. Soc. B304*, 411-424 (1984).
- Mann, S. in *Biomineralization in Lower Plants and Animals* vol. 30 (ed. Leadbetter, B. S. C. & Riding, R.) 39-54 (Systematics Ass., Oxford University Press, 1986).
- Mann, S. *J. Inorg. Biochem.* **28**, 363-370 (1986).
- Weiner, S. *CRC Crit. Rev. Biochem.* **20**, 365-408 (1986).
- Wheeler, A. P. & Sikes, C. S. *Am. Zool.* **24**, 933-944 (1984).
- Addadi, L. & Weiner, S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4110-4114 (1985).
- Mann, S., & Perry, C. C. *Ciba Fdn Symp.* **121**, 40-58 (1986).
- Mann, S., Sparks, N. H. C. & Weiner, S. *Proc. R. Soc. B* (in the press).
- Mann, S., Hannington, J. P. & Williams, R. J. P. *Nature* **324**, 565-567 (1986).
- Schmidt, F. O. *Proc. Am. phil. Soc.* **100**, 476-486 (1956).
- Smith, J. W. *Nature* **219**, 157-158 (1968).
- Miller, A. *Phil. Trans. R. Soc. B304*, 455-477 (1984).
- Fraser, R. D. B., Macrae, T. P., Miller, A. & Suzuki, E. *J. molec. Biol.* **167**, 497-521 (1983).
- Katz, E. P. & Li, S. J. *J. molec. Biol.* **73**, 351-369 (1973).
- Glimcher, M. J., Hodge, A. J. & Schmitt, F. O. *Proc. natn. Acad. Sci. U.S.A.* **43**, 860-867 (1957).
- Degens, E. T., Spencer, D. W. & Parker, R. H. *Comp. Biochem. Physiol.* **20**, 553-579 (1967).
- Hare, P. E. *Science* **139**, 216-217 (1963).
- Weiner, S. *Am. Zool.* **24**, 945-951 (1984).
- Meenakshi, V. R., Hare, P. E. & Wilbur, K. M. *Comp. Biochem. Physiol.* **40B**, 1037-1043 (1971).
- Weiner, S. & Traub, W. *FEBS Lett.* **111**, 311-316 (1980).
- Weiner, S., Talmon, Y. & Traub, W. *Int. J. biol. Macromolecules* **5**, 325-328 (1983).
- Weiner, S. *Phil. Trans. R. Soc. B304*, 425-434 (1984).
- Greenfield, E. M., Wilson, D. C. & Crenshaw, M. A. *Am. Zool.* **24**, 925-932 (1984).
- Sikes, C. S. & Wheeler, A. P. in *Biomineralization and Biological Metal Accumulation* (ed. Westbroek, P. & de Jong, E. W.) 285-289 (Reidel, Amsterdam, 1983).
- Lee, S. L., Veis, A. & Glonek, T. *Biochemistry* **16**, 2971-2979 (1977).
- Crenshaw, M. A. & Ristedt, H. *Biomineralization* **8**, 1-8 (1975).
- Lee, S. L., Glonek, T. & Glimcher, M. J. *Calcif. Tissue Int.* **35**, 815-818 (1983).
- Spector, A. R. & Glimcher, M. J. *Biochim. biophys. Acta* **303**, 360-362 (1973).
- Lee, S. L. & Glimcher, M. J. *Calcif. Tissue Int.* **33**, 385-394 (1981).
- Glimcher, M. J. in *The Chemistry and Biology of Mineralized Connective Tissues* (ed. Veis, A.) 618-673 (Elsevier, New York, 1981).
- Kretsinger, R. K., & Nelson, D. J. *Coordination Chem. Rev.* **18**, 29-124 (1976).
- Weiner, S. *Biochemistry* **22**, 4139-4144 (1983).
- Weiner, S. & Hood, L. *Science* **190**, 987-989 (1975).
- Addadi, L., Moradian, J., Shay, E., Maroudas, N. G. & Weiner, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2732-2736 (1987).
- Mann, S. & Calvert, P. D. *Trends Biotechnol.* (in the press).
- Calvert, P. D. & Mann, S. *J. Mater. Sci.* **5**, 309-314 (1987).
- Landau, E. M., Levanon, M., Leiserowitz, L., Lahav, M. & Sagiv, J. *Nature* **318**, 453-456 (1985).
- Landau, E. M. *et al. Molec. Cryst. Liq. Cryst.* **134**, 323-335 (1986).
- Mechanical Design in Organisms* (eds Wainwright, S. A., Biggs, W. D., Currey, J. D., & Gosline, J. M.) (Princeton University Press, 1982).
- Nakahara, H. in *Biomineralization and Biological Metal Accumulation* (ed. Westbroek, P. & de Jong, E. W.) 225-230 (Reidel, Amsterdam, 1983).

A method for prediction of volcanic eruptions

Barry Voight

Department of Geosciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

The relationship $\dot{\Omega}^{-\alpha}\ddot{\Omega} - A = 0$ describes the behaviour of materials in terminal stages of failure, where Ω is an observable quantity such as strain, and A and α are empirical constants. Drawing on analogies between failure mechanics and eruption processes at volcanoes, Ω is interpreted in terms of conventional geodetic, seismic or geochemical observations. Manipulation of Ω provides a consistent analytical basis for eruption prediction.

THE more successful empirical models for predicting volcanic eruptions probably have a common foundation in first principles or natural laws. Although the linkages may seem obscure, theoretical investigations may establish a consistent framework to aid the understanding of empirical relations. Predictions based on such fundamentals are likely to be more reliable than purely probabilistic attempts based on pattern recognition¹. The purposes of this paper are to: (1) introduce a fundamental relationship describing pre-eruption rate changes of various phenomena; (2) develop equations that couple rate changes of volcanic processes to the time of event occurrence; (3) emphasize the predictive utility of the inverse representation of rate; and (4) apply the predictive method to examples of several specific eruptions.

The basis for the method is the recognition that a fundamental law for failing materials:

$$\dot{\Omega}^{-\alpha}\ddot{\Omega} - A = 0 \quad (1)$$

where A and α are empirical constants and Ω is an observable quantity, suitably describes the behaviour of many quantities preceding a volcanic eruption^{2,3}. The dot refers to differentiation with respect to time. Quantities that might be represented by Ω include strains or displacements for rocks, soils, alloys, metals, plastics, concrete or ice. By encompassing volcanology within the realm of 'rock mechanics', Ω is interpreted in terms of conventional geodetic observations (such as length change, fault slip, strain or angular change), seismic quantities (such as the square root of cumulative energy release) or geochemical observations (such as gas emission rates or chemical ratios) (see Fig. 1).

Volcanic eruptions are invariably preceded and accompanied by rock failure. In volcanic systems, deterministic failure mechanics models may apply to brittle fracture produced in

surrounding rock by magma pressure, or by the pressure of volatiles exsolved from magma at various levels, from deep systems (≥ 5 km) to the surficial domes. Dome rupture, conduit clearing or sector collapse influenced by magmatic or volatile pressures all imply damage and stresses in excess of breaking strength. Earthquakes commonly associated with such damage may be 'tectonic-like' (for example, associated with fault slip induced by magmatic pressure), or of the 'shallow-volcanic' or 'surface-event' varieties^{4,5}. Gas emission rates and chemical variation may be directly influenced by permeability, pressure and pressure-gradient changes associated with brittle fracture^{6,7}, and hence may relate to damage-mechanics models.

Theory

Equation (1), a relationship first noted in experimental landslips caused by monotonic load increase^{8,9}, is perceived as a fundamental physical law governing diverse forms of material failure for conditions of constant stress and temperature^{2,3}. If Ω is taken to be strain, for example, the final stages of failure under steady conditions of an alloy in tension, a rock in compression or a tilting volcanic dome would show a proportionality between the logarithm of creep acceleration and the logarithm of creep velocity. Integrating equation (1) gives expressions for the rate, $\dot{\Omega}$. For $\alpha = 1$, and with $\dot{\Omega} = \dot{\Omega}_0$ at $t = t_0$, we have

$$\dot{\Omega} = \dot{\Omega}_0 e^{A(t-t_0)} \quad (2)$$

For $\alpha < 1$,

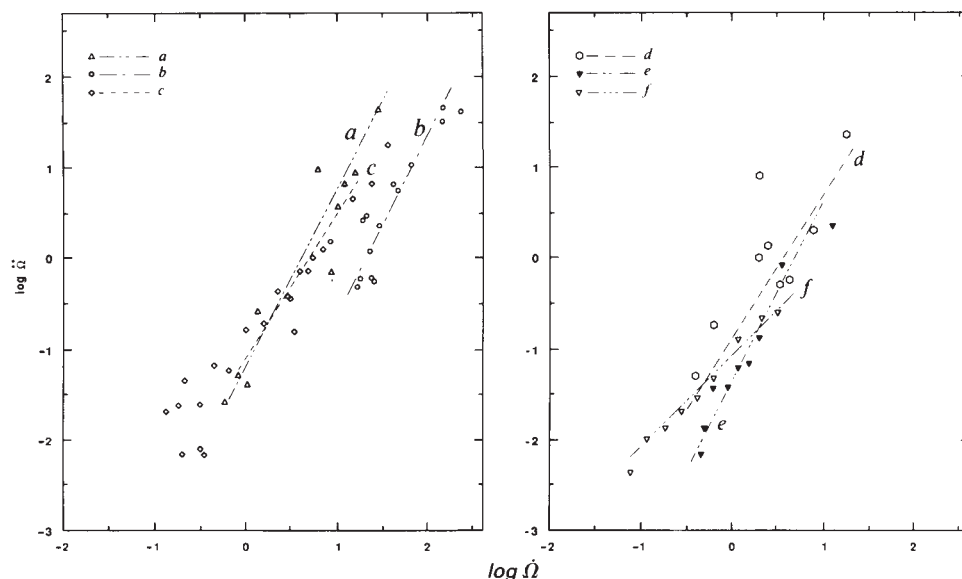
$$\dot{\Omega} = [A(1-\alpha)(t-t_0) + \dot{\Omega}_0^{1-\alpha}]^{1/(1-\alpha)} \quad (3a)$$

or for $\alpha > 1$,

$$\dot{\Omega} = [A(\alpha-1)(t_f-t) + \dot{\Omega}_f^{1-\alpha}]^{1/(1-\alpha)} \quad (3b)$$

where t_f is the time of failure and $\dot{\Omega}_f$ the rate at failure (see Fig. 2).

Fig. 1 Relationship between $\ddot{\Omega}$ and daily rate $\dot{\Omega}$ preceding eruption or failure, for different events and meanings of Ω . *a*, Ω = line length change (cm), Mt St Helens, 1982, $\alpha = 1.96$; *b*, Ω = tilt (μ rad), Mt St Helens, 1982, $\alpha = 2$; *c*, Ω = fault movement (cm), Mt St Helens, 1981, $\alpha \approx 1.4$; *d*, Ω = cumulative seismic strain release ($10^3 \text{ J}^{1/2}$), Bezymyanny, 1960, $\alpha = 1.6$; *e*, Ω = surface movement (cm), Mt Toc, 1963, $\alpha = 1.97$; *f*, Ω = surface movement (cm), Mt Toc, 1960, $\alpha = 1$.



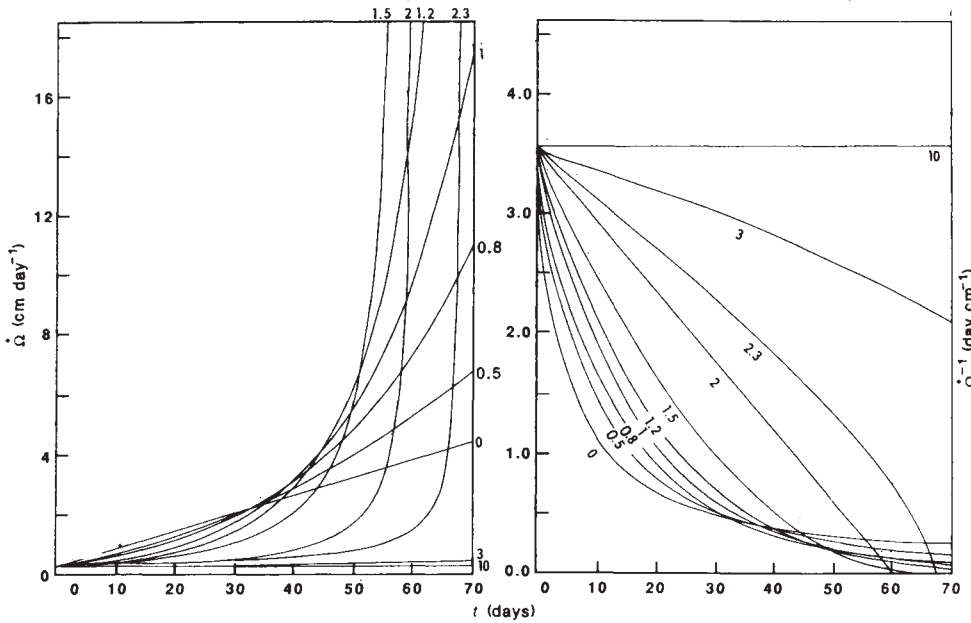


Fig. 2 Rate (left) and inverse rate (right) against time, showing curvature for different α values as indicated. Input simulates example for line length change at dome: $A = 0.059$, $\Omega_0 = 0.28 \text{ cm day}^{-1}$, $t_0 = 0$. Equation (2) for curve $\alpha = 1$; otherwise equation (3a). For 19 March 1982 eruption at Mt St Helens, actual $\alpha \approx 2$.

Expressions for Ω arise from double integration of equation (1). For $\alpha = 1$,

$$\Omega - \Omega_0 = \frac{\dot{\Omega}_0}{A} (e^{A(t-t_0)} - 1) \quad (4)$$

For $\alpha = 2$, and neglecting the $\dot{\Omega}_t$ term,

$$\Omega - \Omega_0 = \frac{1}{A} [\ln(t_f - t_0) - \ln(t_f - t)] \quad (5)$$

For $\alpha > 1$, $\alpha \neq 2$,

$$\begin{aligned} \Omega - \Omega_0 = & \frac{1}{A(\alpha-2)} \\ & \times \{ \dot{\Omega}_0^{2-\alpha} - [A(1-\alpha)(t-t_0) + \dot{\Omega}_0^{1-\alpha}]^{(2-\alpha)/(1-\alpha)} \} \end{aligned} \quad (6a)$$

or

$$\begin{aligned} \Omega - \Omega_0 = & \frac{1}{A(\alpha-2)} \{ [A(\alpha-1)(t_f - t_0) + \dot{\Omega}_f^{1-\alpha}]^{(2-\alpha)/(1-\alpha)} \\ & - [A(\alpha-1)(t_f - t) + \dot{\Omega}_f^{1-\alpha}]^{(2-\alpha)/(1-\alpha)} \} \end{aligned} \quad (6b)$$

By manipulation of equation (3), at any arbitrary $t = t_*$, where $\dot{\Omega} = \dot{\Omega}_*$, the time to failure can be calculated:

$$t_f - t_* = \frac{\dot{\Omega}_*^{1-\alpha} - \dot{\Omega}_f^{1-\alpha}}{A(\alpha-1)} \quad (7)$$

For $\dot{\Omega}_f$ assumed infinite

$$t_f - t_* = \frac{\dot{\Omega}_*^{1-\alpha}}{A(\alpha-1)} \quad (8)$$

gives an upper-bound solution. Neglect of the terminal rate leads to non-conservative forecasts, but the error is often small and may not be important to decision makers.

In the analytical method using equations (7) or (8), values for A and α may be estimated from a plot such as that in Fig. 1, or from solving a set of simultaneous equations like equation (3), using rates observed up to the current time. Least-squares regression provides an estimate of error. The accuracy of the method is determined by the precision and frequency of observations and the regularity and continuity of the observed phenomena. The log rate-log time relation of Fig. 3 is useful for hindcast calculations but requires knowledge of the eruption date and cannot be used for prediction.

The time of eruption t_e is related to the time of failure t_f by $t_e = t_f + f(t)$, where $f(t)$ is an arbitrary function. This simply indicates that there can be a time delay between rupture of a conduit at depth or surficial collapse of an endogenous dome, and the venting of eruption products. Although some volcanoes may require different functions, in many cases a delay interval expressed in hours or days may be sufficient. Under other circumstances t_e may be simply equated to t_f .

The time to failure can also be ascertained graphically using the reciprocal-rate curve. Studies of the perception of curves that grow with time demonstrate that forecast accuracy is aided by inverse representation^{10,11}. As large differences in such sequences occur early rather than late, trends may be recognized early. As used here, the inverse rate

$$\dot{\Omega}^{-1} = [A(1-\alpha)(t-t_0) + \dot{\Omega}_0^{1-\alpha}]^{1/(\alpha-1)} \quad (9)$$

decreases continuously with time, linearly for $\alpha = 2$, upwardly convex for $\alpha > 2$ and concave for $\alpha < 2$ (Fig. 2). Experience suggests that for volcanoes, α is frequently nearly 2. In such cases the inverse form is nearly linear⁹, and graphical extrapolation is straightforward. Failure occurs when the inverse rate value $\dot{\Omega}_f^{-1}$ is obtained; for $\dot{\Omega}_f^{-1}$ neglected, failure is approximated as the point of intersection of the reciprocal rate curve with the time axis.

Both this extrapolation and equations (7) or (8) enable the prediction of specific (if often preliminary) eruption dates, rather than merely broad 'time windows' within which eruptions are considered likely. Predictions are easily updated with new data. In experienced hands, the method provides a way to improve the art of prediction at particular volcanoes, and may allow extrapolation of predictive patterns to volcanoes elsewhere.

The preceding relations apply where loading is relatively steady. For variable stress (σ), equations (2) and (3) may be expanded as follows. For $\alpha = 1$,

$$\dot{\Omega} = \dot{\Omega}_0' (\sigma/\sigma')^m e^{A(t-t_0)} \quad (10)$$

for $\alpha \neq 1$,

$$\dot{\Omega} = \{A(1-\alpha)(t-t_0) + [\dot{\Omega}_0' (\sigma/\sigma')^m]^{1-\alpha}\}^{1/(1-\alpha)} \quad (11a)$$

or for $\alpha > 1$,

$$\dot{\Omega} = \{A(\alpha-1)(t_f - t) + [\dot{\Omega}_f' (\sigma/\sigma')^m]^{1-\alpha}\}^{1/(1-\alpha)} \quad (11b)$$

where $\dot{\Omega}_0'$ and $\dot{\Omega}_f'$ are initial and failure rates associated with a stress σ' , and m is an empirical constant. For variable stress

Fig. 3 Relationship between logarithm of rate and logarithm of time preceding eruption or failure. Negative slopes $n = 1/(\alpha - 1)$ for data are as follows: *a*, line length change, Mt St Helens, 1982, $n = 1.04$; *a'*, same, 3 days before eruption, $n = 0.73$; *b*, tilt, Mt St Helens, 1982, $n = 1.0$; *c*, fault, Mt St Helens, 1981, $n = 1.5$; *d*, cumulative seismic strain release, Bezymyanny, $n = 1.66$; *e*, surface movement, Mt Toc, 1963, $n = 1.03$.

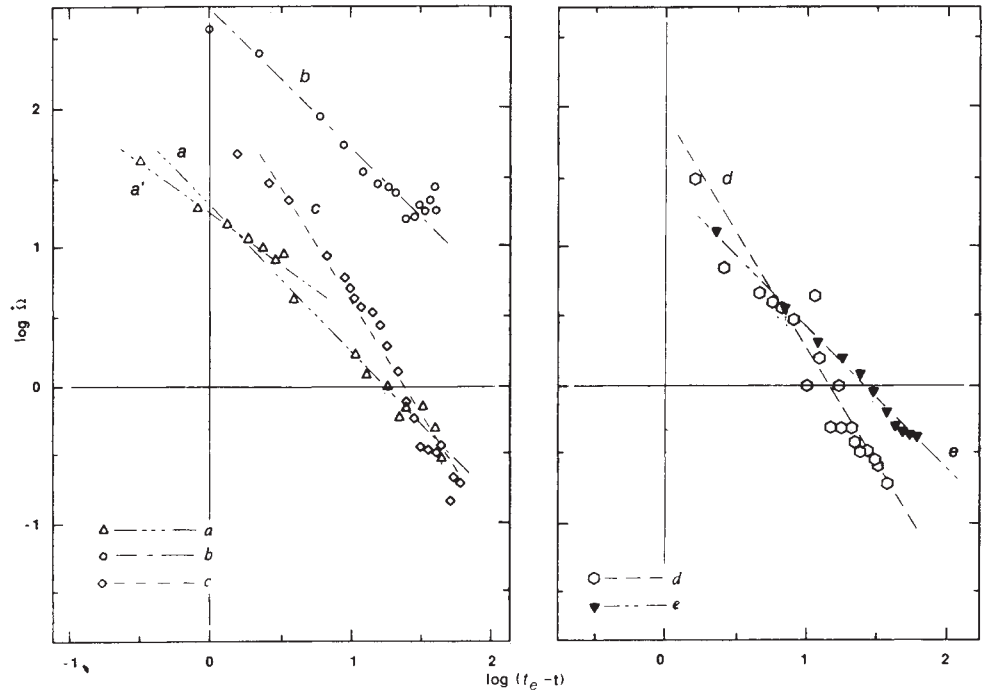
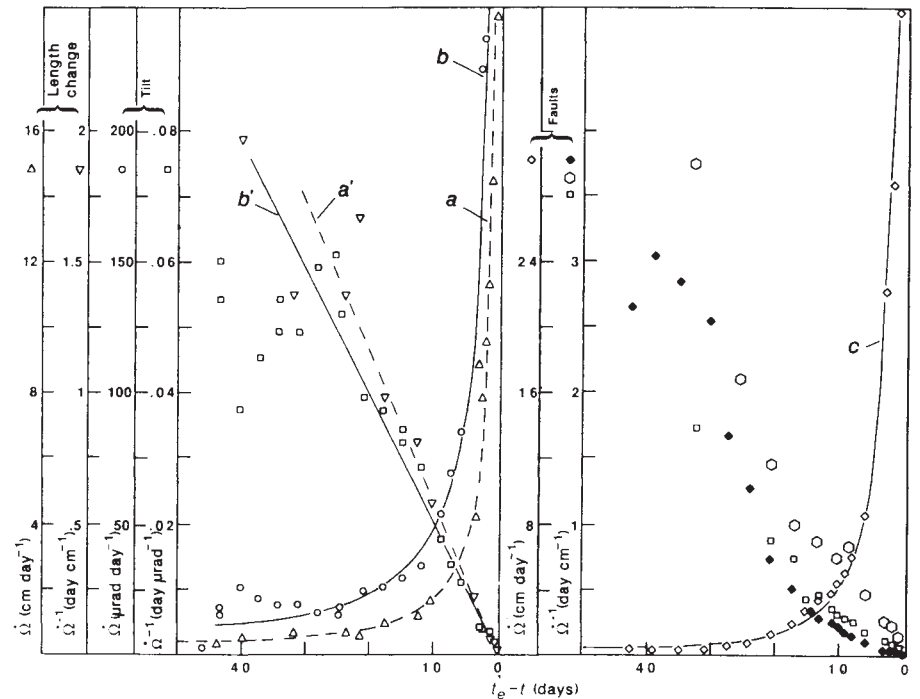


Fig. 4 Mount St Helens. Left, rates and inverse rates against time before the 19 March 1982 eruption. Ω is line length change (curves *a*, *a'*) or tilt (*b*, *b'*), with primes indicating the inverse-rate curves. For both data sets, $\alpha = 2$. Right, Christina 2 fault movement (curve *c*, diamonds) before the September 1981 eruption, $\alpha = 1.4$. Inverse rates for two other faults indicated by squares and hexagons.



histories, failure life is governed by a life fraction rule of the type

$$\sum_{i=1}^n (t_i/t_f^i) = 1, \quad (12)$$

where t_i is the time over which stress σ_i is applied, and t_f^i is the rupture time corresponding to the constant stress σ_i .

Special cases encompassed by equation (1) and related expressions include many independently discovered empirical or quasi-deterministic equations that describe time-deformation relations preceding failure. Such cases, which to a certain extent provide a validation of equation (1), include the Rabotnov-Kachanov equations of materials science¹², the Monkman-Grant equation for deformation of metals and alloys¹³, the general Saito expression¹⁴⁻¹⁶ (developed for soils but applicable to various materials), and exponential idealizations of tertiary creep¹⁶.

For example, the general Saito expression^{15,16} for creep rate $\dot{\epsilon}$,

$$\dot{\epsilon} = E(t_f - t)^{-n} \quad (13)$$

is equivalent to equation (3) when $\Omega = \epsilon$, $E = [A(\alpha - 1)]^{1/(1-\alpha)}$, $n = 1/(\alpha - 1)$, for the special case of negligible $\dot{\Omega}_t^{1-\alpha}$ and $\alpha > 1$. The expression simplifies for $\alpha = 2$ ($n = 1$). Similarly, the Monkman-Grant expression¹³,

$$t_f \dot{\epsilon}_s^q = \text{constant} \quad (14)$$

between time to fracture t_f and minimum creep rate $\dot{\epsilon}_s$, is duplicated with equation (3) by setting $\Omega = \epsilon$, $\dot{\epsilon} = \dot{\epsilon}_s$ at $t = t_0 = 0$, and neglecting $\dot{\Omega}_t^{1-\alpha}$, yielding

$$t_f \dot{\epsilon}_s^{\alpha-1} = 1/A(\alpha - 1) \quad (15)$$

The Monkman-Grant expression is thus a special case of equation (1), for $\dot{\Omega}_t$ large and $\alpha > 1$, where $\alpha = q + 1$. Because

Monkman-Grant-type plots ordinarily employ total rupture life rather than time in tertiary creep, an adjustment in timescale may be necessary for the calculation of equivalent constants.

Applications

In the following, theory is discussed in relation to three eruptions at Mt St Helens, USA, two eruptions at Bezymyanny Volcano, USSR, and a major rockslide at Mt Toc, Italy. The examples have been chosen to illustrate the manipulation of data from diverse geodetic and seismic observations in order to provide guidance for practical applications in eruption forecasting.

Line length changes at Mount St Helens dome. On 12 March 1982 a "relatively long-term prediction" was issued for the Mount St Helens volcano, indicating that "an eruption is likely within the next three weeks"¹. By 15 March, with greatly accelerated deformation, this prediction was updated at 19:00 PST (local time): "An eruption, most likely of the dome building type, will probably begin within 1 to 5 days." At 09:00 PST on 19 March, a "relatively short-term prediction" based on increased seismicity during the day stated that "an eruption would begin soon, probably within 24 hours". The eruption began at 19:27 PST on 19 March (see cover).

Deformation monitoring from the crater floor was critical to the "relatively long-term" predictions. Experience at Mount St Helens has shown that inflation, recognized by contraction of certain measured distances, typically begins weeks before an eruption occurs. Letting Ω = length change, data were evaluated from curve 2 of ref. 1 by graphical differentiation. Slopes were established by coordinate differences for adjacent data points and assigned to time-interval midpoints; log velocity was then plotted against $\log(t_e - t)$, where t_e is the time of eruption (Fig. 3a, a'). The surface deformation conditions here seem appropriate for the assumption $t_e = t_f$. The data are well described by equation (3) or (13), $\Omega = 20(t_e - t)^{-1.04}$, where the rate $\dot{\Omega}$ is in cm per day, time is in days and the exponent n is equal to $1/(\alpha - 1)$. Considering Fig. 3a and a' in detail, a divergence in slope can be seen ~3 days before the eruption, with the final trend better represented by $\dot{\Omega} = 17.4(t_e - t)^{-0.73}$ (curve a'). Comparable information may be gained from the plot of log acceleration against log velocity (Fig. 1a), although the variance is larger owing to the double differentiation necessary for acceleration. The complete data set is well fit by equation (1), $\dot{\Omega} = 0.059\ddot{\Omega}^{1.96}$, an expression equivalent to $\dot{\Omega} = 20(t_e - t)^{-1.04}$.

The effect of α on rate is illustrated in Fig. 2 (left). Although for each curve an increase of rate with time seems to be relatively distinct, beyond any arbitrary point in time, it is difficult to achieve an accurate visual extrapolation of rate increase. The value $\alpha = 1.96$, however, indicates a nearly linear function of inverse rate with time (compare $\alpha = 2$ curve in Fig. 2 (right), and curve a' in Fig. 4). Nearly uniform decrease of the inverse rate allows graphical extrapolation of the time of eruption to be made with reasonable confidence. Alternatively, pre-eruption estimates of α and A may be used to calculate the time of eruption from equation (7) or (8).

With this method, an experienced interpreter might be able to recognize predicted eruption windows earlier than was formerly possible. Specific eruption dates could, under favourable circumstances, be targeted from the "relatively long-term" perspective—which for the 19 March example was >10 days in advance of the eruption. The method is iterative and permits sequential adjustments with new data.

Finally, accumulation of α and A values for various pre-eruptive phenomena allows experience to be quantified and perhaps applied with greater precision in future events. The efficiency of eruption prediction should be enhanced at specific volcanoes, and quantified patterns might be successfully extrapolated to volcanoes elsewhere, even if exact values are not duplicated.

Tilt measurements. Here Ω represents tilt data from Mount St

Helens for the period February–March 1982, for comparison to the distance measurements for the same 19 March eruption of the previous example, and to illustrate the importance of redundancy and cross-correlation of data in eruption prediction. Both radial and tangential tilt at one station (ROA) resumed in mid-January and accelerated fairly smoothly through March¹⁷. A break in trend and some irregularity occurred on both curves in early to mid-February.

In curve b of Fig. 3, the log rate of tilt ($\dot{\Omega}$ is angular velocity, in μ rad per day) is plotted against $\log(t_e - t)$. Here too the data seem to fit reasonably well the simplified form with $\alpha = 2$, less weight being accorded to the disturbed data of mid-February, corresponding to $\log(t_e - t) \approx 1.5$. This representation gives $\dot{\Omega} = 500(t_e - t)^{-1}$, and is equivalent to $\ddot{\Omega} = 0.002\dot{\Omega}^2$, where $\ddot{\Omega}$ is angular acceleration. The latter expression is tested in Fig. 1 (curve b).

Actual tilt rates plotted against time in Fig. 4 (curve b) show a close relation to the $\alpha = 2$ approximation. In this same figure (curve b'), it is not surprising to see the inverse rate $\dot{\Omega}^{-1}$ descending in nearly linear fashion with time from an irregular plateau of high values. This descent seems recognizable ~18 days before the eruption. If the data had been graphically interpreted in this fashion on 12 March, when an eruption was predicted to occur "within three weeks", a narrow window of, say, 15–21 March might have been suggested. For $A \approx 0.002 (\pm 0.0005)$, application of equation (7) or (8) would have suggested initiation of eruption between 17 and 22 March.

Faults activated by a magmatic system. Several data sets can be used to produce independent α values and inverse rate against time curves. These may exhibit different negative slopes and curvatures, but ideally should converge toward a common eruption date. This concept is illustrated by 1981 data on movement of thrust faults outward from the volcanic dome at Mount St Helens^{1,18}. Here Ω represents cumulative contraction of taped distances across a thrust-fault toe. As rates of observed crater-floor deformation were comparable to those preceding earlier eruptions, a "relatively long range prediction" was issued on 26 August, stating that an eruption was likely within a 2-week period starting 2 September¹. A "relatively short-term prediction" and a revision were issued on 6 September, the day of lava extrusion. Swanson *et al.*¹ note that data were available on 3 September for an improved update of the long-term prediction.

Concave-upward curvature of the inverse rate versus time plot for the Christina 2 fault (Fig. 4 solid diamonds) indicates $\alpha < 2$. For $A \approx 0.1$ (Fig. 1c), equation (8) indicates correctly the date of eruption over 10 days in advance to an accuracy of a few days. An increase of α with time is suggested by Fig. 3c, implying an increase in linearity of the inverse-rate curve and increasingly less ambiguous graphical extrapolation to the time of eruption.

The relationships of inverse rate to time for three independent data sets from different faults are compared in Fig. 4. Both convex and concave curvatures are displayed (implying, respectively, $\alpha > 2$ and $\alpha < 2$), but nonetheless the data points converge towards the date of actual eruption, 6 September.

Seismic energy release. Characteristics of the seismic activity of volcanoes include the energy E_s , expressed in joules, and the cumulative strain release $E_s^{1/2}$. It has been recognized that andesitic eruptions may be preceded by a systematic increase in the cumulative strain release^{5,19–21}; continuous seismic monitoring may thus serve as a basis for eruption prediction. In practice, surface events may be distinguished from medium to low-frequency events, and the square root of energy release can be plotted separately for each group. Both types of event appear to increase before eruptions, but at different rates⁵.

As an example²¹, we consider Ω = cumulative strain release in units of $10^3 \text{ J}^{1/2}$ for the Bezymyanny, Kamchatka, eruption of 12 April 1960. The eruption date is indicated by t_e . Graphical differentiation, both from individual data and from the smoothed curve²¹, yields cumulative strain rates $\dot{\Omega}$. The logarithms of these quantities are plotted against $\log(t_e - t)$ in

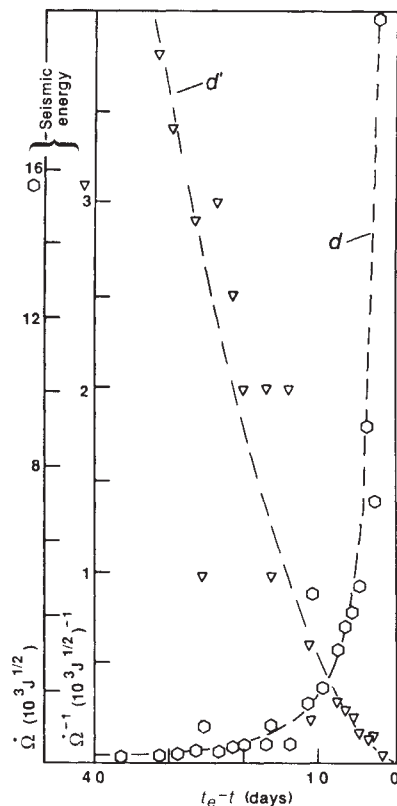


Fig. 5 Bezmyanny. Cumulative seismic strain release rate Ω (curve d) and inverse rate Ω^{-1} (d') against time before 12 April 1960 eruption.

Fig. 3. The distribution of data is linear and well fit by equation (3) or (13), $\Omega = 80(t_e - t)^{-1.66}$. The equivalent expression for equation (1) is $\Omega = 0.12\Omega^{1.6}$ (Fig. 1d).

Actual strain rates are in accord with model curves (Fig. 5), which grow rapidly as eruption impends. The inverse-rate curve is concave upwards but not emphatically so, and linear extrapolation within the 10 days preceding the eruption would have indicated the correct date of eruption to an accuracy of a few days. Thus, equation (1) shows promise of utility, in at least some circumstances, in eruption prediction with seismic quantities²². In some instances an appropriate delay interval of hours to days should separate the time of eruption t_e from the peak of seismic energy release denoted by t_f .

Volcanic debris avalanches and directed-blast volcanism. Anticipation of slide depth, and an estimate of the likelihood of interception of a magmatic-hydrothermal system, are essential for forecasting the possibility of blast volcanism or post-slide plinian activity associated with rockslides and avalanches from an active volcano. For such an association, the problem of predicting explosive volcanism is virtually that of predicting slide failure. This is an instance in which paroxysmal explosivity may be expected at the outbreak of an eruption. Among the most reliable and least ambiguous of observational data for prediction of the time of slope failure are displacements; instrumental extensometer and tilt measurements can provide continuous data²³. As in 1986 at Nevado del Ruiz, Colombia, such data may make possible the forecast that slope collapse is improbable, thus reducing the concerns faced by decision makers considering evacuation and other mitigation procedures²⁴.

Prediction of time to slope failure may be based on velocities or on strains. In the absence of a quantitatively predicted volcanic avalanche, an example of the method is given by the 270 million cubic metre rockslide of Mt Toc, Italy—a slide volume

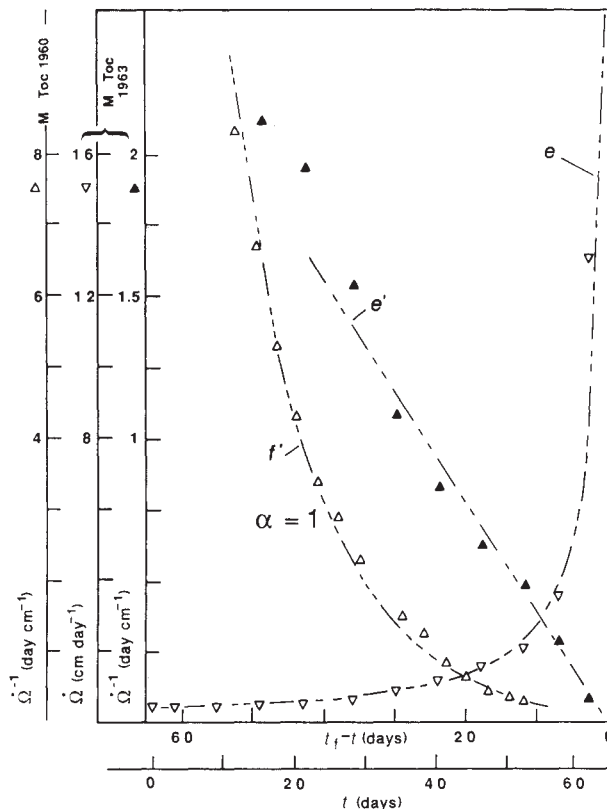


Fig. 6 Mt Toc. Displacement rate Ω (curve e) and inverse-rate Ω^{-1} (curve e') against time before 9 October 1963 slope failure. Exponential inverse-rate against time (curve f') for Fall 1960 movement.

rarely exceeded in historic times²⁵⁻²⁸. The 1963 slide moved a 250-m-thick rock mass 300–400 m horizontally, coming to rest after running up the opposing valley wall. Here, Ω is taken to be horizontal displacement.

A plot of $\log \Omega$ against $\log(t_f - t)$ for movement data over the period 6 August to 9 October 1963 (Fig. 3e), suggest that the data are well expressed by $\Omega = 28.8(t_f - t)^{-1.03}$, or within error limits, $\Omega = 27(t_f - t)^{-1}$. Corresponding acceleration-velocity relationships are $\Omega = 0.040\Omega^{1.97}$, or $\Omega = 0.037\Omega^2$.

The velocity-time plot demonstrates the adequate fit of these expressions to the data (Fig. 6e, e'). With the inverse rate linear for $\alpha \approx 2$, a practical prediction of the failure date could have been made in late September, more than 10 days before the actual failure on 9 October. The high velocity of the slide created a surge of water in a reservoir, overtopping a dam, killing more than 2,000 people, and causing losses totalling 200 million (1964) US dollars²⁵.

A special case is illustrated by an earlier movement phase, in the autumn of 1960, for which α is nearly unity (Figs 1f, 6f'). For $\alpha = 1$ the integral of equation (1) leads to a simple exponential, as confirmed by a plot of $\ln \Omega$ against time, $\Omega = 0.042e^{0.081t}$.

On 18 May 1980, a rockslide-avalanche triggered a volcanic lateral blast at Mount St Helens, Washington^{29,30}. Major lateral displacement of a sector of the cone produced extensional deformation across a kilometre of the north summit area, first observed on 27 March. This suggested that the 'toe' of a potential slide was being compressed, and the possibility of a cubic-kilometre-scale avalanche and associated catastrophic blast³¹ was recognized before the start of geodetic monitoring on 23 April³²⁻³⁴. Monitoring soon indicated deformation rates of ~ 2 m per day on the high north flank of the cone^{29,35}. A close association of earthquake foci with the bulging area led most scientists studying the volcano to conclude correctly in May that a shallow magmatic intrusion was taking place^{32,33}, but before 18 May

there was no scientific consensus on future expectations, and thus no "institutional judgment"³⁶. Among alternative outcomes considered were the emplacement of a volcanic dome or a north-flank slide much smaller than that which occurred on 18 May^{35,37}.

The rate of movement was expected to increase before any actual slide detachment³⁵, but was never observed to occur generally except near the toe. With the available observational data, the method of inverse rate against time could not have been used to predict the occurrence of the 18 May Mount St Helens avalanche; however, a related method seems applicable.

Displacement rates were so large that, if sustained for sufficient time, slope failure was certain. This statement is equivalent to a failure criterion based on strain or total displacement, implying a relationship between rupture life and strain, such as equation (14) or (15). Based on equation (14), Saito and Uezawa³⁸ plotted data from compression tests on different soils, many of volcanic derivation, and reported

$$\log t_f = 2.33 - 0.916 \log \dot{\epsilon}_s \pm 0.59 \quad (16)$$

and a simplified form assuming an exponent of unity,

$$t_f = 214 \dot{\epsilon}_s^{-1} (\pm 0.59) \quad (17)$$

where t_f is rupture life in minutes from the inception of loading and $\dot{\epsilon}_s$ is minimum creep rate in units of 10^{-4} min^{-1} . Although based on laboratory experiments, equation (16) is considered applicable to full-scale slope failure^{38,39}. If loading at Mount St Helens occurred at $\sim 2 \text{ m}$ per day over a 2 km base length; $\dot{\epsilon}_s \approx 0.001 \text{ day}^{-1} \approx 0.007 \times 10^{-4} \text{ min}^{-1}$. From equation (16), rupture is predicted within the range 4–55 days, as reckoned from very late in March⁴⁰. The simplified form gives confidence limits of 5 to 82 days. These estimates enclose the 18 May failure of Mount St Helens, for which t_f is ~ 53 days.

The 30 March 1956 event at Bezmyanny, Kamchatka, provides the only other available data set for an avalanche and directed blast⁴¹. A 230-m horizontal displacement of the cone was detected in conjunction with dome growth, probably occurring within the period November 1955 to February 1956, suggesting an average rate of $\sim 2.5 \text{ m}$ per day. The minimum rate was less than this, and the base length was somewhat less than 2 km , so here too $\dot{\epsilon}_s$ seems to be $\sim 0.001 \text{ day}^{-1}$ —about the same as at Mount St Helens. Predicted time to failure is thus ~ 5 –82 days,

not quite encompassing the actual time to failure at Bezmyanny of ~ 120 days.

As a predictor, this empirical relationship is only expected to have order-of-magnitude accuracy. Its importance is to focus attention on the inevitability of sector failure of volcanic cones within months if magmatic pressure induces sustained high rates of lateral displacement.

An improved relationship could consider such factors as scale dependency, geometry of loading, lithology, stress and movement history, equivalent viscosity of the magma, intrusion rate, and chamber or conduit geometry. The scale effect, for example, would imply that larger masses may be able to absorb larger strain before failure, consistent with the compliancy of a rock mass in comparison to compliancy of intact small specimens tested in the laboratory. Based on the two cases considered, times of failure for cone sectors of order 1 km^3 may be better approximated by $t_f = 800 \dot{\epsilon}_s^{-1}$, with an expected accuracy of perhaps 0.6 log units [that is, $\log t_f = \log (800 \dot{\epsilon}_s^{-1}) \pm 0.6$]. This relationship would predict $t_f \approx 80$ days (range 20–320 days) for $\dot{\epsilon}_s \approx 0.001 \text{ day}^{-1}$.

The surface motions of rock adjacent to magma could overwhelm the nearly elastic-range local motions preceding fracture growth at depth in limited parts of the system⁴⁰, and acceleration of the latter could go undetected if not specifically searched for. The possibility exists in hindsight for Mount St Helens that increased emphasis on geodetic surface monitoring in the vicinity of the toe might have revealed accelerations pertinent to evaluating stability²³. Quasi-steady motion of rock slopes above an immobile toe should be recognized as concentrating stress near the toe; if stress concentration becomes sufficiently large, collapse of the toe could precipitate catastrophic release of the entire slope and trigger off a directed blast.

This work has benefited from my experience with the US Geological Survey, particularly at Cascades Volcano Observatory, from US A.I.D.-supported tasks concerning volcano hazard mitigation in Colombia and the Philippines, and from invited travel to National Research Centers for Disaster Prevention in Japan. I thank R. Janda, R. Decker, C. Newhall, D. Swanson, T. Nakamura, S. Okuda, S. Oyagi, M. Calvache, N. Banks, S. Malone, D. Dzurisin, T. Casadevall, A. Swan, R. Dionne, J. Dutton and P. Luckie, among others, for stimulation of thought, hospitality or assistance; J. Ewart for checking the mathematics; and R. L. Christiansen for a constructive review.

Received 20 July 1987; accepted 18 January 1988.

- Swanson, D. A. *et al. Science* **221**, 1369–1376 (1983).
- Voight, B. *Eos* **68**, 1551 (1987).
- Voight, B. *Eos* **68**, 1286 (1987).
- Weaver, C. S. *et al. Science* **221**, 1391–1394 (1983).
- Malone, S. *et al. Science* **221**, 1376–1378 (1983).
- Tazieff, H. in *Forecasting Volcanic Events* (eds Tazieff, H. & Sabroux, J.-C.) Ch. 21 (Elsevier, Amsterdam, 1983).
- Casadevall, T. *et al. Science* **221**, 1383–1385 (1983).
- Fukuzono, T. & Terashima, H. *Rep. natn. Res. Center Disaster Prev.* **29**, 103–122 (1982).
- Fukuzono, T. in *Proc. 4th Natl. Conf. Field Workshop on Landslides* 145–150 (Disaster Research Institute, Kyoto, 1985).
- Timmers, H. & Wagenaar, W. A. *Percept. Psychophys.* **21**, 558–562 (1977).
- Wagenaar, W. A. in *Environmental Assessment of Socioeconomic Systems* (eds Burkhardt, D. F. & Ittelson, W. H.) (Plenum, New York, 1978).
- Leckie, F. A. in *Mechanics of Engineering Materials* (eds Desai, C. S. & Gallagher, R. H.) Ch. 21 (Wiley, New York, 1984).
- Monkman, F. C. & Grant, N. J. *Proc. Am. Soc. Test. Mater.* **56**, 593–605 (1956).
- Saito, M. *Proc. 7th int. Conf. Soil Mech. Found. Engng* **2**, 677–683 (1969).
- Yamaguchi, S. *Rep. Sogu Doboku Lab. No. 14* (Tokyo, 1978).
- Varnes, D. J. *Proc. 7th Southeast Asian geotech. Conf.* **2**, 107–130 (1983).
- Dzurisin, D. *Science* **221**, 1381–1383 (1983).
- Chadwick, W. W. Jr *et al. Science* **221**, 1378–1380 (1983).
- Tokarev, P. I. *Eruptions and Seismic Regime of Klyuchevskaya Group Volcanoes*, (Nauka, Moscow, 1966) (in Russian).

- Tokarev, P. I. *Dokl. Akad. Nauk SSSR* **199**, 422–425 (1971).
- Tokarev, P. I. in *Forecasting Volcanic Events* (eds Tazieff, H. & Sabroux, J.-C.) Ch. 19 (Elsevier, Amsterdam, 1983).
- Wu, F. T. & Thomson, L. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **12**, 167–173 (1975).
- Voight, B. *Eos* **68**, 456 (1987).
- Voight, B. *et al. in Proc. 6th Congr. int. Soc. Rock Mechanics* (eds Horget, G. & Vongpaisal, S.) pp. 275–279 (A. A. Balkema, Rotterdam, 1987).
- Muller, L. *Rock Mech. engng Geol.* **2**, 148–212 (1964).
- Selli, R. & Trevisan, L. *Annali del Museo Geologico di Bologna* **32**, 8–104 (1964).
- Hendron, A. J. Jr & Patton, F. D. *Geotechnical Laboratory tech. Rep. No. GL-85-5* (U.S. Army Engineer exp. Sta., Vicksburg, 1985).
- Voight, B. & Faust, C. *Géotechnique* **32**, 43–54 (1982).
- Voight, B. *et al. Géotechnique* **33**, 243–273 (1983).
- Voight, B. *et al. Prof. Pap. U.S. geol. Surv.* **1250**, 347–377 (1981).
- Voight, B. Slope stability hazards, Mt St Helens Volcano, Washington. *Tech. Rep. U.S. geol. Surv.* File report, Cascades Volcano Observatory, Washington (1980).
- Decker, R. W. *Prof. Pap. U.S. geol. Surv.* **1250**, 815–820 (1981).
- Decker, R. W. *A. Rev. Earth planet. Sci.* **14**, 267–291 (1986).
- West, S. *Science News* **118**, 399 (1980).
- Lipman, P. W. *et al. Prof. Pap. U.S. geol. Surv.* **1250**, 143–156 (1981).
- Wesson, R. L. *Yb. U.S. geol. Surv.*, FY 1980 (1982).
- Miller, D. *et al. Prof. Pap. U.S. geol. Surv.* **1250**, 789–802 (1981).
- Saito, M. & Uezawa, H. *Proc. 5th int. Conf. Soil Mech.* **1**, 315–318 (1961).
- Saito, M. *Proc. 6th int. Conf. Soil Mech. Found. Engng* **2**, 537–541 (1965).
- Voight, B. in *Proc. int. Conf. Landslides*. State-of-art (oral) report, Toronto (1984).
- Gorshkov, G. S. *Bull. volcan.* **20**, 77–109 (1959).

Perinatal lethal osteogenesis imperfecta in transgenic mice bearing an engineered mutant pro- $\alpha 1(I)$ collagen gene

Alex Stacey*, John Bateman†, Ted Choi*, Tom Mascara†, William Cole† & Rudolf Jaenisch*

* Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Massachusetts 02142, USA

† Department of Paediatrics, University of Melbourne, Parkville, Victoria 3052, Australia

Substitutions of single glycine residues of $\alpha 1(I)$ collagen have previously been associated with the inherited disease osteogenesis imperfecta type II. Transgenic mice bearing a mutant $\alpha 1(I)$ collagen gene into which specific glycine substitutions have been engineered show a dominant lethal phenotype characteristic of the human disease, and demonstrate that as little as 10% mutant gene expression can disrupt normal collagen function.

TYPE I collagen is a major protein of the extracellular matrix. Each molecule is composed of two $\alpha 1(I)$ and one $\alpha 2(I)$ chains wound together as a triple helix. The α -chains are ~1,000 amino acids long, and consist in the triple helical region of repeated triplets of the amino acids Gly-X-Y, where X and Y are frequently proline residues. The α -chains are synthesized as pre-pro- α -chains which undergo post-translational modification, including enzymatic hydroxylation of specific prolyl and lysyl residues and glycosylation of some hydroxyllysyl residues. Triple helices form by association of pro- α -chains at their carboxyl termini with the progressive folding of the helix towards the amino termini. The occurrence of glycine at every third position is critical for triple helix formation which prevents the further modification of the α -chains^{1,2}. The helical procollagen molecules are secreted into the extracellular space, where the amino-terminal and carboxyl-terminal propeptides are cleaved and fibrils are formed.

Malfunction of type I collagen has been observed in many forms of the inherited disorder osteogenesis imperfecta (OI) (refs 3–5). The most severe form (OI type II) is an autosomal dominant lethal disease, resulting in death *in utero* or shortly after birth. Affected infants are small, have a fragile, poorly mineralized skeleton and commonly have collagen reduced in amount and an over-modification of type I collagen^{6–12}. It has been proposed that over-modification of type I procollagen results from structural defects in the α -chains that interfere with their ability to form a triple helix, and that the consequent delay in helix propagation leads to excessive exposure of the α -chains to modifying enzymes^{8–11}. The molecular basis of the dominant defect in four cases of OI type II has been identified as a deletion^{13–17} or as point mutations in the pro- $\alpha 1(I)$ collagen gene *COL1A1*. Each of the three point mutations resulted in the substitution of a glycine residue either by arginine at position 391 (ref. 18) or by cysteine at position 988 (refs 9, 19) or 748 (ref. 20).

The genetic determinants important for normal collagen structure and function are as yet defined only by such rare, sporadic mutations in humans. A more powerful approach for studying the molecular pathology of collagen is in the generation of animal models with defined mutations in the collagen genes. Here we have introduced substitutions for glycine analogous to those in human OI patients into the mouse pro- $\alpha 1(I)$ collagen gene by site-directed mutagenesis. The functional consequences

of the expression of these mutant genes were examined both in cell culture and in transgenic mice to determine whether a dominant OI phenotype could be reproduced.

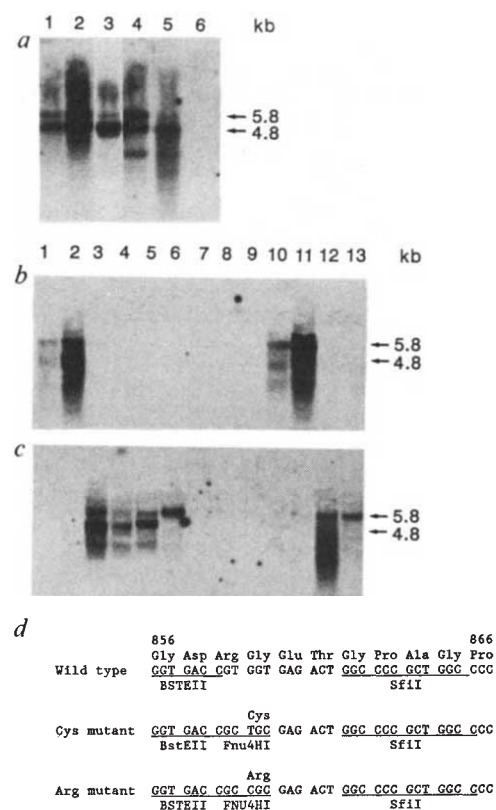
Mutagenesis of the *COL1A1* gene

Two mutations were produced in a mouse *COL1A1* genomic clone, both of which resulted in the substitution of a glycine residue at position 859 of the $\alpha 1(I)$ chain by either a cysteine or an arginine. Mutagenesis of the cloned gene was carried out by replacement of the wild-type DNA region between *Bst*EII and *Sfi*I restricted sites at positions 2,623 and 2,641 (ref. 21) by double-stranded synthetic oligodeoxynucleotides. A *Fnu*4HI restriction site was created at the site of each mutation as a marker. The nucleotide and amino-acid sequences of the wild type and of each mutation are given in Fig. 1d. The final constructs contained the complete coding region of the gene linked to one of four promoters: one-kilobase (kb), 2.5-kb and 3.7-kb fragments 5' of the mouse *COL1A* coding region, or the Moloney murine leukaemia virus long terminal repeat promoter region. The mutant constructs were transfected into either NIH3T3 fibroblasts or into Mov13 homozygous fibroblasts (designated M13, ref. 23), which do not express the pro- $\alpha 1(I)$ collagen chain^{22,23}. Southern analysis showed the presence of multiple copies of the mutant gene in transfected cell clones (data not shown).

Mutant collagen RNA expression in transfected M13 cells was detected by hybridization of a Northern blot with a pro- $\alpha 1(I)$ collagen complementary DNA clone²⁴. Figure 1a shows that NIH3T3 cells (lane 5), as well as M13 cells transfected with the Gly-Cys mutant construct (lanes 1–4), synthesized the normal 5.8-kb and 4.8-kb pro- $\alpha 1(I)$ collagen RNA species which were absent in the parental M13 cell line (lane 6). To specifically detect mutant pro- $\alpha 1(I)$ collagen RNA in the presence of normal RNA, each mutant construct carried a 21-base pair (bp) insert (derived from linker and polylinker remnants) at the *Xba*I site in the 5' untranslated region²². Two synthetic oligomers homologous to the sequence of the inserts in the 2.5-kb and the 3.7-kb promoter constructs and short flanking regions were used to specifically detect mutant pro- $\alpha 1(I)$ collagen RNA by Northern blot analysis. Figure 1b shows RNA transcribed from the 2.5-kb promoter Gly-Cys construct in M13 cells (lanes 1 and 2) and in NIH3T3 cells (lanes 10 and 11). Figure 1c shows RNA from the 3.7-kb promoter Gly-Arg construct in Mov13 cells

Fig. 1 Northern blot analysis of RNA from transfected M13 cell clones (the M13 cell line has been derived from homozygous Mov13 embryo²³) hybridized to pHUC, a human pro- $\alpha 1(I)$ collagen cDNA clone²⁴. Lanes 1–4, M13 Gly-Cys clones LTR-1, LTR-2, 2.5-1, 2.5-2; lane 5, NIH3T3; lane 6, untransfected M13. The transfected cell clones are designated by the mutation (Gly-Cys or Gly-Arg) and the type of promoter (LTR or 1.0, 2.5, 3.7 kb of 5' flanking sequences) and the cell clone number. In addition to the normal $\alpha 1(I)$ collagen RNA species, smaller RNA bands are observed that have been noted previously²³ and arise from truncated collagen genes in the transfected cell clones. **b** and **c**, Northern blot analysis of RNA from transfected M13 and NIH3T3 cell clones showing the specific detection of RNA transcribed by mutant constructs. Panels **b** and **c** are autoradiographs of the same filter; **b** shows hybridization to a 33-base oligomer (ATGCTAGAGTCGACCGGGGATCCTCTAGACCC) specific to the 2.5 kb promoter Gly-Cys construct, **c** shows hybridization to a 31-base oligomer (TGTCTAGAGTCGAGGGGGGATCCTCTAGACC) specific to the 3.7-kb promoter Gly-Arg construct. Lanes 1, 2, M13 Gly-Cys clones 2.5-1, 2.5-2; lanes 3–6, M13 Gly-Arg clones 3.7-3, 3.7-7, 2.7-8, 3.7-10; lanes 7, 8, untransfected M13; lane 9, NIH3T3; lanes 10, 11, NIH3T3 Gly-Cys clones 2.5-4, 2.5-6; lanes 12, 13, NIH3T3 Gly-Arg clones 3.7-2, 3.7-14.

Methods. A fragment between the unique *NheI* and *XhoI* sites at the gene²¹ was sub-cloned, mutated and inserted into the mouse pro- $\alpha 1(I)$ collagen genomic clones derived from the cosmid clone 10D (ref. 23). Four constructs which carried different promoter regions were used: these were 1-kb, 2.5-kb and 3.7-kb fragments of the mouse pro- $\alpha 1(I)$ -collagen promoter region extending from the *XbaI* site immediately 5' of the start of translation to upstream *PvuII*, *HindIII* and *XbaI* sites respectively²², and a *Sau3A* fragment of the Moloney murine leukaemia virus long terminal repeat linked to the gene at the same *XbaI* site. Mutant constructs were cotransfected into the Mov13 homozygous embryo fibroblast cell line M13 (ref. 23) and NIH3T3 cells with a selectable marker pAG60, which contains the neomycin phosphotransferase gene conferring resistance to the drug G418 (ref. 28), at a molar ratio of 7:1. Transfection was carried out as described²⁹. Following transfection, total cellular RNA was prepared³⁰. RNA samples (10 μ g) were denatured with glyoxal, separated on a 1% agarose gel and transferred to Gene Screen plus nylon membrane (NEN). Single-stranded oligonucleotide probes were 5' end-labelled using [γ -³²P]ATP and T4 polynucleotide kinase, pHUC was ³²P-labelled by random primer extension, using the Klenow fragment of *E. coli* DNA polymerase I. Hybridization of oligonucleotide probes was in 50% formamide, 5 $\times$ SSPE (750 mM NaCl, phosphate buffered), 1% SDS, 0.1 mg ml⁻¹ sonicated salmon sperm DNA, 1 $\times$ Denhardt's solution, at 37 °C. The final wash was 0.5 $\times$ SSPE (75 mM NaCl), 0.1% SDS at 45 °C. Hybridization to pHUC was as above but at 45 °C. The final wash was in 0.1 $\times$ SSPE (15 mM NaCl), 0.1% SDS at 45 °C. NIH3T3 cells were selected in 1 mg ml⁻¹ G418 (Gibco) and M13 cells in 0.3 mg ml⁻¹ G418. The names of individual clones refer to the original cell line, the mutation present, the promoter and the clone number. **d**, Nucleotide and amino acid sequence of mouse pro- $\alpha 1(I)$ collagen between amino acid residues 856 and 866 of the α -helical region. The substitution of glycine for cysteine and arginine in each mutant is indicated. Underlined are the *BstEII* and *SfiI* restriction sites between which synthetic oligomers were inserted, and also the *Fnu4HI* sites created in each mutant.



(lanes 3–6) and in NIH3T3 cells (lanes 12 and 13). The mutant genes express the characteristic 4.8-kb and 5.8-kb pro- $\alpha 1(I)$ collagen RNA species, although their relative abundance varies between clones, as previously observed²³. The data in Fig. 1 show that most clones produced full-length mutant collagen RNA in amounts comparable to the endogenous RNA in 3T3 cells. The extent of expression from constructs with different promoters was similar.

Characterization of mutant collagen

The production of stable collagen protein was assessed by SDS-PAGE of [³H]proline-labelled collagen after limited digestion with pepsin. Pepsin digestion removed the N- and C-terminal procollagen propeptide domains but did not degrade the stable triple-helical α -chain regions of the native collagen molecule. Hence, pepsin digestion can simplify the gel pattern by converting all the procollagens to collagen, and can indicate whether a stable collagen triple helix has formed.

The expression of the pepsin-stable collagen in the transfected M13 cells is shown in Fig. 2a. In cell lines expressing the pro- $\alpha 1(I)$ Gly-Cys gene (Fig. 2a, lanes 1–4), a significant proportion of the pepsin-stable collagen was present as a component with an electrophoretic migration comparable to that reported for the disulphide-linked $\alpha 1(I)$ chain dimers observed in cases of lethal perinatal OI with cysteine mutations⁹. This band was reducible to give another band migrating slightly more slowly than an $\alpha 1(I)$ chain (data not shown). The gel profile of the reduced chain after cyanogen bromide cleavage was similar to that of an $\alpha 1(I)$ chain (Fig. 2c; ref. 18), except that each of the main bands migrated more slowly, as would result from over-modification⁸. We conclude that the reduced chain consists of

an over-modified $\alpha 1(I)$ chain. In the transfected cells (Fig. 2c, lanes 3 and 4) the carboxy-terminal $\alpha 1(I)$ CB6 peptide was only apparent after reduction. These data were consistent with a glycine to cysteine mutation at residue 859 in the $\alpha 1(I)$ chain, which results in disulphide bonding between the $\alpha 1(I)$ CB6 regions of these chains.

In contrast, in the Gly-Arg mutant chain (Fig. 2a, lanes 6–8) the predominant collagen product migrated like the $\alpha 1(I)$ chain. CNBr-peptide mapping confirmed that this was an $\alpha 1(I)$ chain, which contained a more basic charged form of the $\alpha 1(I)$ CB6 peptide, as expected (data not shown). A similar charge change has been noted in the CB8 peptide as a result of a Gly-Arg mutation at residue 391 of the $\alpha 1(I)$ chain in a patient with lethal perinatal OI (ref. 18).

The M13 cells, although deficient in endogenous $\alpha 1(I)$ expression, synthesize $\alpha 2(I)$ chains which can form pepsin-stable collagen molecules with $\alpha 1(I)$ chains produced from the transfected wild-type pro- $\alpha 1$ collagen gene (Fig. 2a, lanes 5, 9). The Gly-Cys mutant chains associate with the endogenous $\alpha 2(I)$ chains to form pepsin-stable helical molecules (lane 1, 2). Some cell clones expressed only low amounts of $\alpha 2(I)$ messenger RNA and therefore $\alpha 1(I)$ trimers were predominantly formed (lanes 3, 4, 6, 7, 8). The total amount of pepsin-stable collagen observed in a given transfectant frequently did not correlate with the steady-state levels of mutant RNA. Thus, clone M13 Gly-Cys 2.5-1 synthesized high concentrations of $\alpha 1$ chains (Fig. 2a, lane 1) but only moderate amounts of the Gly-Cys procollagen mutant RNA (Fig. 1a, lane 3; b, lane 1), whereas clones M13 Gly-Cys LTR-2 or 2.5-2 produced very little pepsin-stable collagen (Fig. 2a, lanes 4, 2) but high amounts of mutant RNA (Fig. 1a, lanes 2, 4; b, lane 2). We do not know whether this is

Fig. 2 *a* and *b*, Protein analysis of transfected cell lines. [^3H]proline-labelled pepsin-resistant collagen produced by mouse fibroblast cultures. *a*, Mov13 cells transfected with the Gly to Cys mutant gene (lane 1, M13 2.5-1; lane 2, M13 2.5-2; lane 3, M13 LTR-1; lane 4, LTR-2) and the Gly to Arg mutant (lane 6, M13 3.7-5; lane 7, M13 3.7-9; lane 8, M13 3.7-10) and with the normal mouse pro- $\alpha 1(\text{I})$ gene (lanes 5 and 9). All samples were from the cell layer fraction and were analysed unreduced. *b*, NIH3T3 cells transfected with the Gly to Cys mutant gene (lanes 1 and 5, 3T3 1.0-1; lanes 2 and 6, 3T3 1.0-3; lanes 3 and 7, 3T3 1.0-4) and the Gly to Arg mutant (lanes 4 and 8). The cell layer (lanes 1-4) and the medium samples (lanes 5-8) were analysed separately. The migration position of $\alpha 1(\text{I})$, $\alpha 2(\text{I})$, $\alpha 1(\text{V})$, $\alpha 2(\text{V})$ and the mutant disulphide-linked $\alpha 1(\text{I})$ dimer (*) are shown. The arrowheads (*b*) indicate the slowly migrating $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ chains. *c*, CNBr peptide analysis of pepsin-stable collagen produced by Mov13 cells transfected with the normal pro- $\alpha 1(\text{I})$ gene (lanes 1 and 2) and with the Gly to Cys mutant M13 2.5-1 (lanes 3 and 4). Samples in lanes 2 and 4 were reduced with 10 mM dithiothreitol before electrophoresis. The migration position of the major $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ CNBr peptides and the mutant $\alpha 1(\text{I})$ CB6 (*) are indicated. Samples were digested in 70% formic acid containing 50 mg ml $^{-1}$ CNBr (ref. 31) before lyophilization and analysis on 12.5% polyacrylamide gels 32 .

Methods. Mov13 and NIH3T3 cell cultures were set up at equal cell densities in 150 cm 2 dishes and grown until 1-2 days after visual confluency in Dulbecco's Modified Eagle's Medium (DMEM) containing 10% (v/v) fetal calf serum. Collagen labelling and analysis was as described previously 8,10 . Pre-incubation for 4 h was in 20 ml DMEM containing 10% dialysed fetal calf serum and 0.15 mM sodium ascorbate and 0.1 m β -aminopropionitrile. L-[5- ^3H]Proline (27 Ci mmol $^{-1}$, Amersham) was added (10 $\mu\text{Ci ml}^{-1}$) and labelling was continued for 18 h. Culture medium and cell layer fractions were separated. Procollagen and collagen were precipitated with 25% saturated ammonium sulphate and the precipitates collected and dissolved in 0.05 M Tris/HCl, pH 7.5, 0.15 M NaCl. Portions were prepared for electrophoretic analysis by digestion with 100 $\mu\text{g ml}^{-1}$ of pepsin in 0.5 M acetic acid for 16 h at 4°C. Samples were quantitatively loaded and electrophoresis performed on 5% (w/v) polyacrylamide gels containing 2 M urea. Radioactive bands were detected by fluorography.

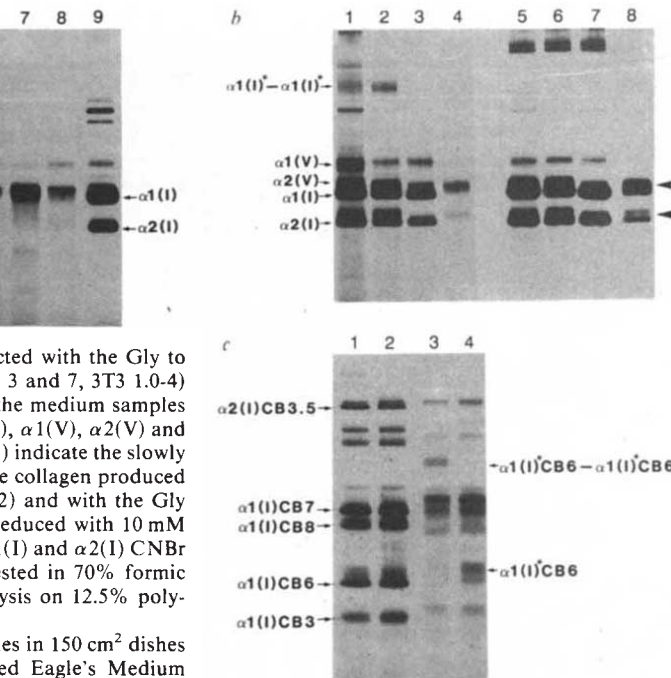
due to altered proteolytic susceptibility of the mutant gene products, to differences in translational efficiencies, or to clonal variation.

The molecular defects in lethal OI have only been analysed so far in individuals heterozygous for the mutation. To mimic this situation we transfected the mutant genes into 3T3 cells which produce the endogenous $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ chains (Fig. 2*b*). The level of pepsin-stable mutant collagen was different in the different cell clones. Clearly defined doublets of both the $\alpha 1(\text{I})$ and the $\alpha 2(\text{I})$ bands in Gly-Cys clones 3T3 1.0-3 and 1.0-4, as well as $\alpha 1(\text{I})$ disulphide-linked dimers (lanes 1 and 2), indicated significant synthesis of mutant $\alpha 1(\text{I})$. Cells transfected with the Gly-Arg mutant (Fig. 2*b*, lanes 4 and 8) also synthesized normal, as well as slow-migrating, forms of both the $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ chains. Similarly, slow-migrating forms of both $\alpha 1$ and $\alpha 2$ chains have been seen in cells from patients with perinatal lethal OI. The retarded migration of these chains is due to increased lysine hydroxylation and glycosylation $^{8-12}$.

All clones secreted mutant collagen into the medium (Fig. 2*b*), but the amount secreted varied between the different cell lines. Impaired secretion of the mutant collagen has been observed in some cases of lethal perinatal OI (refs 8, 12). Another common consequence of glycine mutations in human OI is reduced thermal stability of the triple helix 9,20 , leading to functional impairment of the collagen molecule. We have noted a similar reduced thermal stability of the mutant collagen triple helix produced by the transfected NIH3T3 cells (data not shown).

Generation of lethal OI type II phenotype

These results indicate that cultured cells expressing the mutant constructs produced type I collagen with the same biochemical characteristics as mutant collagens produced by cells from patients with perinatal lethal OI. To determine whether the mutant constructs would generate a dominant phenotype similar to the human disease, DNA from the 3.7 kb-promoter Gly-Cys mutant cosmid was microinjected into fertilized eggs which were



allowed to complete development *in utero*. Three of the 7 foster mothers delivered naturally and their offspring were examined within an hour after birth. Offspring from the other 4 mothers were delivered by caesarian section and examined immediately. DNA from fetuses and surviving animals was assayed for the presence of the transgene by hybridization with pUC13 DNA on slot blots. Table 1 (top) shows that none of the 32 mice surviving after normal birth or caesarian section was positive for the transgene. In contrast, 7 out of 9 fetuses which died shortly before or after delivery carried between 0.25 and 6 copies of the mutant gene (Table 1, bottom). Lethality was not observed when either the wild-type mouse or human *COL1A1* gene was microinjected into mouse zygotes (Table 1, control DNA). So far, four phenotypically normal transgenic lines have been developed which express the introduced wild-type *COL1A* genes (unpublished results). Our observations therefore indicate that the mutant gene exerts a dominant lethal effect.

Figure 3*a* shows 3 littermates from a total litter of 13, delivered by caesarian section, two of which carried the transgene (numbers 25 and 28). Fetus 25 (lower) had died *in utero* shortly before delivery and fetus 28 (left) lived for ~20 min. Fetus 26 (right) was phenotypically normal and did not carry the transgene. The two abnormal embryos were unusually flabby and pliable in that their limbs could be flexed in all directions with no resistance, and the cranium of both was unusually soft (Fig. 3*b*). Radiographs (Fig. 3*c*) showed the skeletons of both transgenic fetuses to be markedly underdeveloped with short wavy ribs, short and broad long bones, and general poor mineralization, similar to that seen in human OI type II. Of the other litters examined, 7 fetuses died around birth, of which five carried the mutant gene. Four of the transgenic fetuses (numbers 1, 15, 22 and 29) had no obvious phenotypic abnormalities, but autopsy showed wavy ribs in two of them (1 and 15).

Pro- $\alpha 1(\text{I})$ collagen RNA and collagen protein were analysed in an attempt to correlate expression of the mutant gene with the observed phenotypes. Expression of the mutant and endogenous pro- $\alpha 1(\text{I})$ collagen genes was distinguished and

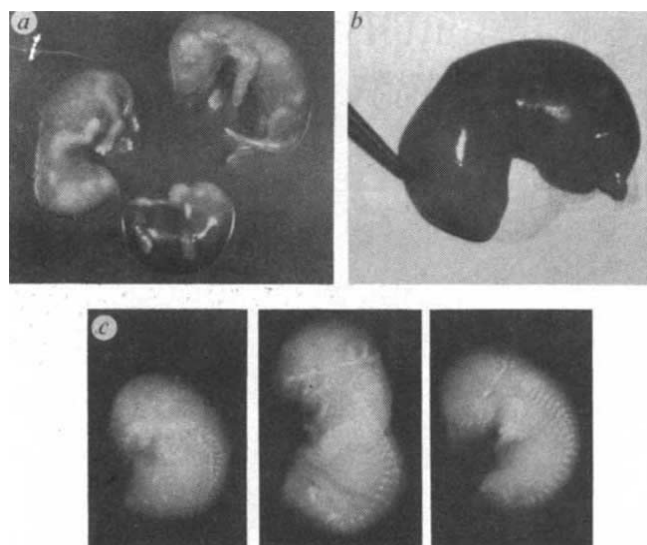


Fig. 3 Phenotype of transgenic newborns. *a*, Photographs of transgenic embryo 25 (lower) and embryo 28 (left) and normal embryo 26 (right). Embryo 25 died shortly before delivery. *b*, Embryo 25: Slight pressure was applied to the head with forceps to demonstrate deficiency of bone formation in skull. *c*, X-ray embryo 25 (left), 26 (middle) and 28 (right). Radiographs were made by 5 s exposure at 40 kV.

quantified by an S_1 -nuclease-protection assay. A 600-bp *Aval* fragment, extending from a position 24 bases 3' of the transcription start site into the first intron, was isolated from the mutant constructs. It was expected that mutant transcripts would protect a 192-base fragment and that endogenous transcripts would protect a 95-base fragment (see Fig. 4b). Figure 4a shows the results of the S_1 -nuclease-protection assay. As expected, the control lanes show that there was no DNA fragment protection by M13 cell RNA (lane 8), but in M13 cells transfected with a mutant construct the expected 192-base fragment was protected (lane 9), and the 95-base fragment in NIH3T3 cells (lane 11), and both fragments in NIH3T3 cells transfected with a mutant construct, were protected (lane 10). Each of the 5 transgenic fetuses from which RNA could be isolated showed protection of the 95- and 192-base fragments, indicating the presence of mutant collagen RNA (lanes 1–5), whereas non-transgenic animals showed only the 95-base fragment (lanes 6 and 7). The relative proportions of mutant and normal RNA were measured by optical densitometry of the autoradiograph, and the results are summarized in Table 1 (bottom). The data show that between 4% and 30% of the total pro- $\alpha 1(I)$ collagen RNA was transcribed from the mutant gene.

Collagen was extracted by pepsin digestion from fetal transgenic mouse skin and analysed by electrophoresis (Fig. 5). The total amount of normal and mutant collagen was quantified and the results summarized in Table 1. A striking reduction of the total collagen I content proportional to the level of mutant gene expression was seen in all transgenic fetuses. Thus, fetuses producing as little as 10% mutant RNA (number 15) had only 46% type I collagen when compared with normal controls. A reduction to 20% was seen in fetus 28, which produced 30% mutant RNA, and the collagen content of fetus 25 was reduced even more. So production of relatively small amounts of mutant procollagen leads to a dramatically reduced type I collagen content. In addition to the reduction of type I collagen in the transgenic fetuses, type III collagen was also reduced, but to a lesser extent (data not shown). Similar results have been observed in the dermis of OI type II patients¹². Most important, our results show that the severity of the pathological phenotype directly correlates with the extent of mutant gene expression, in that fetuses most seriously affected expressed the mutant gene to the greatest extent.

The $\alpha 1(I)$ collagen chains from transgenic fetuses migrated on gels slightly more slowly than normal (Fig. 5a). This was probably a result of loading different amounts of protein on to the gel tracks, because it was not seen when the gels were loaded with equal amounts of type I collagen (Fig. 5b). Under those conditions a small amount of the mutant $\alpha 1(I)$ chain dimer was detected by silver-staining (lane 2) and had an increased mobility after delayed reduction (lane 3). The disulphide-linked dimer was present in all transgenic fetuses in amounts roughly proportional to the level of mutant RNA (data not shown). A slight upward trailing of the α -chains from transgenic fetuses was seen on Coomassie staining of the gel in Fig. 5b, suggesting increased lysine modification (compare Fig. 2b). Due to the small amounts of mutant collagen available, however, it was not possible to quantitate the extent of lysine modification.

Conclusions

Our aim was to generate a mouse-cell culture and animal model of human perinatal lethal osteogenesis imperfecta which has been shown in three cases to result from point mutations of a glycine to a cysteine or arginine residue^{9,18–20}. We introduced identical substitutions at position 859 of the $\alpha 1(I)$ chain by site-directed mutagenesis of a mouse genomic $\alpha 1(I)$ clone. Cells transfected with the mutant genes synthesized collagen $\alpha 1(I)$ and $\alpha 2(I)$ chains having a decreased electrophoretic mobility and hence an increased post-translational lysine modification like the chains produced by cells from human patients. The

Table 1 Generation of transgenic fetuses and mutant gene expression

Generation of transgenic fetuses					
DNA injected	Animals born/ eggs injected	No. of mothers	Delivery	Total no.	No. with transgene
Mutant gene (Gly-Cys)	41/200	3	Spontaneous	Alive 11	0
				Dead 3	2
		4	Caesarian section	Alive 21	0
				Dead 6	5
Control genes	47/220	9	Spontaneous	Alive 47	7
Expression of mutant gene					Type I collagen content (% control)
Animal no.	Copies of mutant gene	% Mutant RNA			
1	6	17			26
5	0.25	ND			ND
15	1	10			46
22	1	4			98
25	4	ND			15
28	6	30			20
29	3	9			89

FVB zygotes were microinjected with $10 \mu\text{g ml}^{-1}$ of the 3.7-kb collagen promoter Gly-Cys mutant cosmid DNA linearized by *Sall* digestion. Control DNA consisted of either a cosmid of the human or the mouse wild-type *COL1A1* gene²³. The injected embryos were transferred to pseudopregnant foster mothers. The embryos were either born spontaneously or delivered by caesarian section. Transgenic fetuses were identified by slot blot hybridization. Mutant gene copy number was estimated after binding genomic DNA to Gene Screen plus membrane filters by slot blotting, hybridization of duplicate filters to pUC13 to detect the vector portion of the mutant transgene, and to the clone p10-2 (ref. 27) to allow quantitation of DNA. Relative proportions of mutant and endogenous $\alpha 1(I)$ procollagen RNA (extracted from total fetus) were determined by optical densitometry of the S_1 protection analysis autoradiograph shown in Fig. 4a. Type I collagen in fetal dermis was quantitated by densitometry of the data in Fig. 5. The values are expressed as % of type I collagen in non-transgenic littermates. ND, Not determined. Only part of embryo number 5 was available for genotyping; embryo 25 had died *in utero* shortly before caesarian delivery and no high molecular weight RNA could be extracted.

Fig. 4 a, Specific detection of endogenous and exogenous pro- $\alpha 1(I)$ collagen gene expression in transgenic mice by S1 nuclease protection assay. Lanes 1–5, RNA from transgenic animals numbers 1, 15, 22, 28 and 29; lanes 6 and 7 are non-transgenic animals, numbers 6 and 16; lane 8, Mov13 homozygous cell line M13; lane 9, M13 Gly-Arg 3.7-7 transfected cell line; lane 10, NIH3T3 Gly-Arg 3.7-2 transfected cell line; lane 11, untransfected NIH3T3. Cell lines transfected with the 3.7-kb promoter Gly-Arg construct, rather than the Gly-Cys construct, were used for convenience, but identical fragments were protected in each case. **b**, Scheme of the S1 assay used to distinguish between mutant and endogenous pro- $\alpha 1(I)$ collagen transcripts. The top line represents the first exon of the mouse pro- $\alpha 1(I)$ collagen gene. The ^{32}P -labelled *Ava*I site and the 21-bp insert present only in the mutant constructs are shown. The probe used was a 600-bp *Ava*I fragment (between the site indicated and a second site in the first intron) from the 3.7 kb promoter Gly-Cys construct. RNA from the exogenous gene constructs protects the *Ava*I fragment over a length of 192 bases from the *Ava*I site to the start of the first intron, as indicated by the middle line. Endogenous RNA which lacks the 21-bp insert protects the probe to a point 3 bases within the insert, resulting in a fragment 95 bases long, as shown by the lower line.

Methods. S1 nuclease protection analysis was carried out by established procedures³³. The *Ava*I fragment was labelled by end-filling with [α - ^{32}P] dCTP and the Klenow fragment of *E. coli* DNA polymerase I. The DNA fragment (~20 ng) was hybridized with 5 μg total RNA from each transfectant at 50 °C for 16 h. S1 nuclease digestion was with 200 units of enzyme (PL Biochemicals) in 300 μl 250 mM NaCl, 10 mM Na Acetate, 1.5 mM ZnSO_4 , 20 mg ml^{-1} sonicated salmon sperm DNA at pH 4.6, 37 °C for 1 h. After phenol extraction and ethanol precipitation, the products were analysed on 7 M urea/6% polyacrylamide gel and visualized by autoradiography.

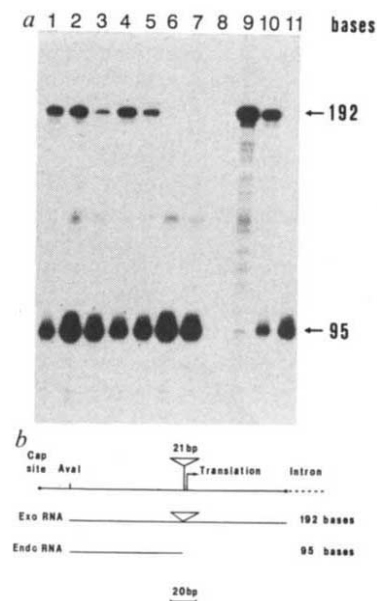
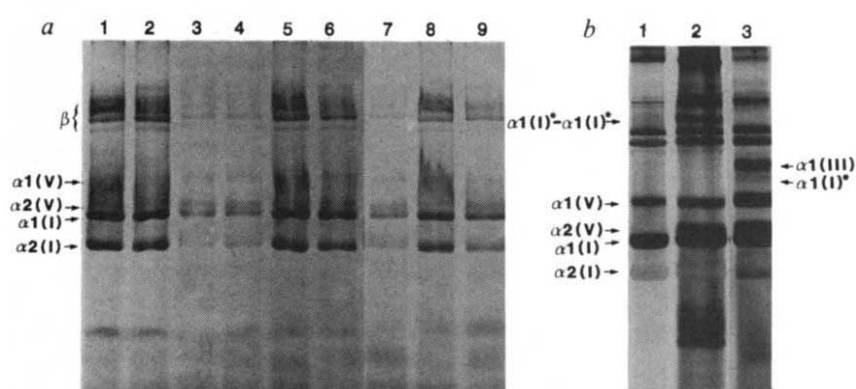


Fig. 5 a, SDS-PAGE of collagen extracted by pepsin from embryo dermis. Lanes 2, 5 and 8, collagen from fetuses which did not contain the transgene; lane 1, fetus 22; lane 3, fetus 25; lane 4, fetus 28; lane 6, fetus 29; lane 7, fetus 1; lane 9, fetus 15. Gel lanes were quantitatively loaded from pepsin extracts of tissue samples of equal dry weight. Samples were not reduced before electrophoresis and gels were stained with Coomassie blue. **b**, Electrophoresis of samples loaded to equal amounts of type I collagen in each lane. Samples in lanes 1 and 2 were not reduced and lane 3 was analysed after delayed reduction. Bands were identified by staining with silver nitrate. Lane 1, collagen from a fetus that did not contain the transgene; lanes 2 and 3, fetus 1. The electrophoretic migration positions of type I collagen β -components, $\alpha 1(I)$ and $\alpha 2(I)$ chains, type III $\alpha 1(III)$ chains and type V collagen $\alpha 1(V)$ and $\alpha 2(V)$ chains are indicated. The disulphide-linked mutant $\alpha 1(I)$ chain dimer (*) and the migration position of the mutant after delayed reduction (*) are also shown.

Methods. Specimens of embryo skin were diced, defatted with chloroform/methanol (ratio 2:1), and lyophilized to constant dry weight. Dermal collagen was then extracted twice by incubation with 100 $\mu\text{g ml}^{-1}$ pepsin in 0.5 M acetic acid at an enzyme to substrate ratio of 1:10 for 24 h at 4 °C. Collagen was precipitated from the clarified extracts by addition of NaCl to a final concentration of 1.5 M. Samples were then quantitatively loaded on to 5% (w/v) polyacrylamide gels for electrophoresis. Protein bands were stained with Coomassie blue³². Samples had equal amounts of collagen in each gel lane and were analysed either with or without delayed reduction with 10% β -mercaptoethanol³⁴; gels were stained with silver nitrate³⁵.



over-modification of lysine found in lethal OI is thought to be a direct consequence of mutations that disrupt the repeating Gly-X-Y amino-acid sequence triplet, leading to disturbance of triple helix propagation^{8,10,11}. Our results demonstrating over-modification of all major $\alpha 1(I)$ and $\alpha 2(I)$ CNBr peptides support this model, as the glycine mutations in the $\alpha 1(I)$ CB6 peptide would result in the exposure of all helical regions of the $\alpha 1(I)$ chains to excessive post-translational modification. The reduced thermal stability of the type I collagen helix was further evidence for its biochemical similarity in the transfected cell cultures and human OI.

Expression of the mutant gene in transgenic mice resulted in a dramatic dominant lethal phenotype, which resembled clinically and biochemically the lethal perinatal phenotype of human patients^{4,8,9,12,13,18}. The extent of mutant gene expression resulted in a correlation of the reduced type I collagen content with the severity of the phenotype. It has been proposed that molecules which contain a mutant chain and have a lower melting temperature are therefore more susceptible to intra- and extracellular degradation. This, in conjunction with poor secretion of the mutant collagen, was postulated to cause the low level of mutant collagen in the dermis²⁵. This model does

not, however, explain why as little as 10% mutant chains should result in such reduced amounts of normal type I, and to a lesser extent type III collagen, leading to a perinatal lethal phenotype. Possibly the intracellular accumulation of the mutant gene product induces degradation of not only abnormal, but also normal, collagen molecules, resulting in the reduced collagen content. In addition, even a small fraction of mutant chains could interfere with correct collagen alignment, fibril formation, or interactions with other components of the extracellular matrix, and result in functional abnormalities. A detailed analysis of helix stability, secretion, procollagen processing and fibril formation in cell lines and transgenic animals carrying mutant collagen genes should lead to a clearer understanding of both normal collagen biosynthesis as well as of the complex phenotype of lethal OI in human patients.

This study constitutes the first example of the generation of an animal model of a human genetic disease by introducing precisely defined alterations into the normal gene and then transferring the mutant gene into transgenic animals. Because the substitution of Gly results in a dominant phenotype, the consequences of the mutation can be tested in normal animals carrying the wild-type alleles. The availability of the Mov13

mouse strain will also permit the analysis of recessive collagen mutations: a mutant gene carried in a transgenic mouse could be crossed into the Mov13 mouse strain, thus eliminating the normal pro- $\alpha 1(I)$ genes and revealing the phenotype of the recessive mutation.

The approach of introducing dominant mutations into animals by gene transfer may have wide applications for the study of disease mechanisms and of protein function. It is likely that mutations in critical regions of proteins which form multimeric structures will have dominant effects when introduced into tissue-culture cells or transgenic animals²⁶. This should allow

the genetic analysis of synthesis and assembly of other collagens, or of cytoskeleton proteins such as actins or intermediate filaments, even without the availability of mutations in the endogenous genes.

We thank Charles Bragdon for help with radiography, Angelika Schnieke and Peter Byers for constructive criticism, and Anton Berns for deriving control transgenic mice carrying the human *COL1A1* gene. This work was supported by Public Health Services grants from the NIH to R.J. and grants from the National Health and Medical Research Council of Australia and the Research Foundation, Royal Children's Hospital.

Received 4 November 1987; accepted 13 January 1988.

- Oikarinen, A., Anttinen, M. & Kivirikko, K. I. *Biochem. J.* **160**, 639–645 (1976).
- Oikarinen, A., Anttinen, M. & Kivirikko, K. I. *Biochem. J.* **166**, 357–362 (1977).
- Prockop, D. J. & Kivirikko, K. I. *New Engl. J. Med.* **311**, 376–386 (1984).
- Byers, P. H. & Bonadio, J. in *Metabolic and Genetic Diseases in Pediatrics* (eds Lloyd, J. & Scriver, C. R.) 56–91 (Butterworth, London, 1985).
- Cheah, K. S. E. *Biochem. J.* **229**, 287–303 (1987).
- Trelstad, R. L., Rubin, D. & Gross, J. *Lab. Invest.* **36**, 501–508 (1977).
- Kirsch, E. *et al. Eur. J. clin. Invest.* **11**, 39–47 (1981).
- Bateman, J. F., Mascara, T., Chan, D. & Cole, W. G. *Biochem. J.* **217**, 103–115 (1984).
- Steinmann, B. *et al. J. biol. Chem.* **259**, 11129–11138 (1984).
- Bonadio, J. F., Holbrook, K. A., Gelinas, R. E., Jacob, J. & Byers, P. H. *J. biol. Chem.* **260**, 1734–1742 (1985).
- Bonadio, J. F. & Byers, P. H. *Nature* **316**, 363–366 (1985).
- Bateman, J. F., Chan, D., Mascara, T., Rogers, J. G. & Cole, W. G. *Biochem. J.* **240**, 669–708 (1986).
- Barsh, G. S. & Byers, P. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5142–5146 (1981).
- Williams, C. J. & Prockop, D. J. *J. biol. Chem.* **258**, 5915–5921 (1983).
- Chu, M.-L. *et al. Nature* **304**, 78–80 (1983).
- Barsh, G. S., Roush, C. L., Bonadio, J., Byers, P. H. & Gelinas, R. E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2870–2874 (1985).
- Chu, M.-L., Gargiulo, V., Williams, C. J. & F. Ramirez, J. *J. biol. Chem.* **260**, 691–694 (1985).

- Bateman, J. F., Chan, D., Walker, I. D., Rogers, J. G. & Cole, W. G. *J. biol. Chem.* **262**, 7021–7027 (1987).
- Cohn, D. H., Byers, P. H., Steinmann, B. & Gelinas, R. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6045–6047 (1986).
- Vogel, B. E., Minov, R. R., Freund, M. & Prockop, D. J. *J. biol. Chem.* (in the press).
- Monson, J. M., Friedman, J. & McCarthy, B. *Molec. cell. Biol.* **2**, 1362–1371 (1982).
- Schnieke, A., Harbers, K. & Jaenisch, R. *Nature* **304**, 315–320 (1983).
- Schnieke, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 764–768 (1987).
- Stacey, A., Mulligan, R. & Jaenisch, R. *J. Virol.* **61**, 2549–2554 (1987).
- Prockop, D. J. *Am. J. hum. Genet.* **36**, 499–505 (1984).
- Herskowitz, I. *Nature* **329**, 219–222 (1987).
- Harbers, K., Schnieke, A., Stuhlmann, H. & Jaenisch, R. *Nucleic Acids Res.* **10**, 2521–2537 (1982).
- Colbere-Garapin, F., Horodniceanu, F., Kourilsky, P. & Garapin, A.-C. *J. molec. Biol.* **150**, 1–14 (1981).
- Graham, R. & Van der Eb, A. *Virology* **52**, 456–457 (1973).
- Auffray, C. & Rougeon, F. *J. Biochem.* **107**, 303–314 (1980).
- Scott, P. G. & Veis, A. *Connect. Tissue Res.* **4**, 107–116 (1976).
- Cole, W. G., Chan, D., Hickey, A. J. & Wilcken, D. E. L. *Biochem. J.* **219**, 451–460 (1984).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
- Sykes, B., Puddle, B., Francis, M. & Smith, R. *Biochem. biophys. Res. Commun.* **72**, 1472–1480 (1976).
- Schleicher, M. & Waterson, M. *Analyt. Biochem.* **131**, 312–317 (1983).

LETTERS TO NATURE

Is the G18.95–1.1 radio source a shell-type supernova remnant with a central object?

A. R. Patnaik, T. Velusamy & V. R. Venugopal

Radio Astronomy Centre, Tata Institute of Fundamental Research, PO Box 8, Udhagamandalam 643001, India

Fürst *et al.*¹ have reported that the nonthermal galactic radio source G18.95–1.1 shows arc-like structures pointing towards the radio peak near its centre, and that its morphology does not conform to any known class of supernova remnants (SNRs). They propose an accreting binary origin for the source, following the suggestion by Helfand and Becker² for the galactic nonthermal radio sources G5.3–1.0 and G357.7–0.1, which show filamentary structure and a compact feature along the axis of symmetry near the brightest edge of the extended emission^{3,4}. But recently these two sources have been shown to be SNRs^{5,6}. The compact feature in G537–0.1 is identified with a compact HII region, and there is no direct evidence for a binary in any of these sources. Although de Kool and van den Heuvel⁷ have modelled these sources as accreting binary systems, the details of conversion of energy into radio filamentary structure are not apparent. In view of the suggested peculiar morphology of G18.95–1.1, high-resolution (96×36 arc s) low-frequency (327 MHz) observations were made using the Ooty Synthesis Radio Telescope (OSRT)⁸. This high-resolution map of G18.95–1.1 is consistent with the conventional shell-type morphology around a central object.

Observations were carried out for about 8 h on each of three days between 6 May and 21 May 1985. A new antenna was later added to the OSRT to provide the short spacings needed for

mapping large angular size structures. New observations were carried out in April 1987. The shortest spacing in our observations is ~ 60 wavelengths, corresponding to an angular size of ~ 60 arc min. The data were calibrated from the source 1830–211 assuming a flux density of 12.4 Jy at 327 MHz. The calibrated visibilities were processed using the Astronomical Image Processing System (AIPS) software package. The clean map was restored with gaussian beam of 96×36 arc s at zero position angle (PA) Fig. 1. The r.m.s. noise in the map is 12 mJy per beam. To compare with the 4.7 GHz map of Fürst *et al.*¹ we also smoothed the map to a resolution of 2.4×2.4 arc min (Fig. 2). The integrated flux density of G18.95–1.1 at 327 MHz is 57.5 ± 5.0 Jy, consistent with an extrapolation of the spectrum from higher frequencies¹.

Whereas the map of Fürst *et al.*¹ exhibits arc-like features, our low frequency map obtained with higher resolution (Fig. 1) reveals an extended shell structure with a bridge of emission connecting the shell at $PA \approx -20^\circ$. The radio peaks near the central emission seen at 4.75 GHz are resolved. The peak brightness occurs on the western shell which is resolved into bright spots and filaments. In Fig. 2, the shell is nearly complete except for a gap along north-eastern boundary which coincides with a minimum in the 4.75 GHz map. It is interesting to note that the strongest emission in the shell is seen over the western part, in the direction of the galactic plane, and such asymmetries are expected due to the decrease in ambient magnetic field, gas and density of relativistic electrons away from the galactic plane.

The overall extent of the source in Fig. 2 (~ 30 arc min) is consistent with that at 4.75 GHz. The differences in the two maps can be attributed to the different observing methods used, a single telescope at 4.75 GHz and a synthesis telescope at 327 MHz. In the synthesis observations, structures larger than 60 arc min are absent. Further, the reconstruction of extended emission (> 15 arc min) is limited by inadequate sampling in the $u-v$ plane with OSRT. We have thus not attempted to derive

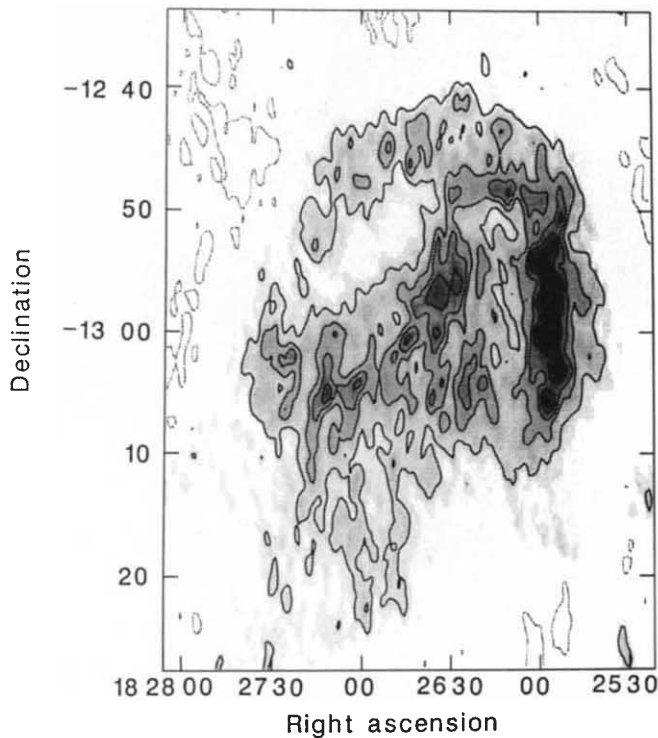


Fig. 1 327 MHz CLEAN map of G18.95-1.1 with a resolution of 96×36 arcs at zero position angle. The grey-scale map is superposed on the contour map. The peak flux density in the map is 334 mJy per beam. The contours are plotted in multiples of 40 mJy per beam.

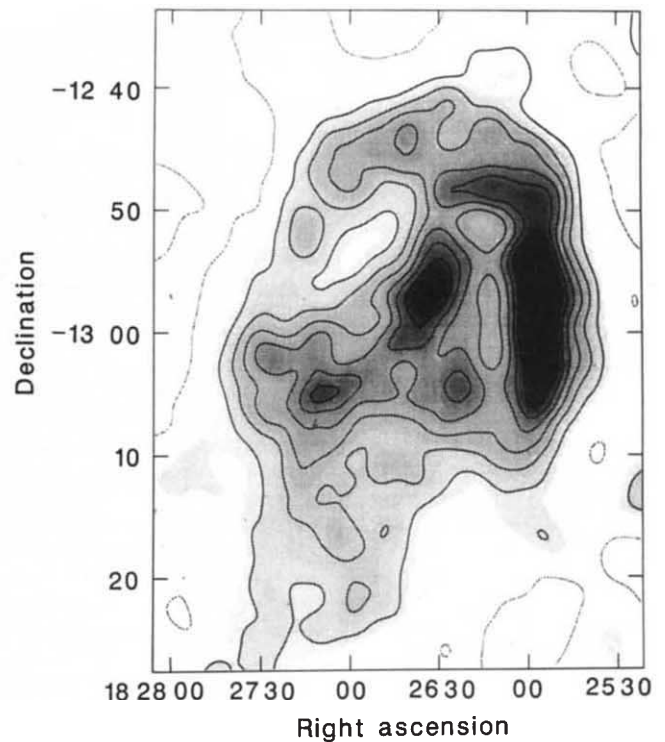


Fig. 2 327 MHz CLEAN map of G18.95-1.1, smoothed to a circular gaussian beam 2.4 arc min in diameter. The grey-scale map is superposed on the contour map. The peak flux density in the map is 1.266 Jy per beam. The contours are plotted in multiples of 120 mJy per beam.

a detailed distribution of the spectral index between 327 MHz and 4.75 GHz over the source. However, a comparison of the 327 MHz map (Fig. 2) smoothed to the resolution of 4.75 GHz map reveals some spectral index variation. For instance, at 327 MHz the brightest emission is marked over the western shell while at 4.75 GHz it is within the shell, near the centre indicating a flatter radio spectrum in the central region. The spectral index α (0.327-4.75 GHz) (defined by flux $S_\nu \approx \nu^\alpha$) over the shell is ~ -0.47 , consistent with the integrated flux density spectral index of -0.4 , between 1.4 and 10.8 GHz, reported by Fürst *et al.*¹. The spectral index for the shell region is close to the mean value for the shell-type SNRs ($\alpha \approx -0.5$). The shell structure, clearly seen in the 327 MHz map, the nonthermal character of emission, and the radio spectral index of the source lead to the conclusion that G18.95-1.1 is a shell-type supernova remnant. Assuming a mean source diameter of 25 arc min and a flux density of 40 Jy at 1 GHz, the distance to G18.95-1.1 is estimated to be about 4 kpc from the Σ - D relation⁹. The age of the SNR is estimated to be $\sim 6,000$ years.

The nature of the emission at the centre of the shell is not clear. Within the shell, the strongest emission at 327 MHz is coincident with the secondary peak in the 4.75 GHz map. This feature is close to the geometric centre of the SNR and may be associated with the remnant; it is clearly nonthermal, with spectral index α (0.327-4.75 GHz) ≈ -0.36 . Some polarization is seen at 4.75 GHz. The strongest emission at 4.75 GHz is displaced

from the centre of the shell and no corresponding peak is observed at 327 MHz. The spectral index is α (0.327-4.75 GHz) ≈ -0.2 , so this region may be thermal. There is no noticeable polarization in the 4.75 GHz map, which supports the interpretation that the feature is an HII region, perhaps unrelated to the SNR.

Although the nature of the central source is difficult to infer from the present data it is reasonable to conclude from the shell-type nature of G18.95-1.1 that the central emission may be associated with the stellar remnant of the supernova. As both the morphology and the radio spectrum of the extended emission conform to shell-type SNRs, there is no need to invoke an accreting binary origin, as was proposed by Fürst *et al.*,¹ but the central region may still contain a binary as the stellar remnant. The characteristics of G18.95-1.1 as revealed by the OSRT map—the shell-type morphology, the emission peak at the centre of the shell and the bridges connecting it to the shell—show marked resemblance to the W50-SS433 system¹⁰ and G109.1-1.1 (CTB109)¹¹. It would be useful to have high resolution multifrequency radio observations and also to search for optical and X-ray emission from this object.

The data were analysed at the National Image Processing Facility for Astronomy (NIPFA) at the Radio Astronomy Centre, Ootacamund. This facility is supported by the Indian Department of Science and Technology.

Received 10 September; accepted 2 December 1987.

1. Fürst, E., Reich, W., Reich, P., Sofue, Y. & Handa, T. *Nature* **314**, 720-721 (1985).
2. Helfand, D. J. & Becker, R. H. *Nature* **313**, 118-119 (1985).
3. Becker, R. H. & Helfand, D. J. *Nature* **313**, 115-118 (1985).
4. Shaver, P. A., Salter, C. J., Patnaik, A. R., van Gorkom, J. H. & Hunt, G. C. *Nature* **313**, 113-115 (1985).
5. Caswell, J. L. *et al. Mon. Not. R. astr. Soc.* **225**, 329-334 (1987).

6. Shaver, P. A. *et al. Astr. Astrophys.* **147**, L23-L24 (1985).
7. de Kool, M. & van den Heuvel, E. P. J. *Nature* **317**, 599-600 (1985).
8. Swarup, G. *J. Astrophys. Astr.* **5**, 139-148 (1984).
9. Milne, D. K. *Aust. J. Phys.* **32**, 83-92 (1979).
10. Downes, A. J. B., Pauls, T. & Salter, C. J. *Mon. Not. R. astr. Soc.* **218**, 393-407 (1986).
11. Hughes, V. A., Harten, R. H., Costain, C. H., Nelson, L. A. & Viner, M. R. *Astrophys. J.* **283**, 147-153 (1984).

Bulk superconductivity at 120 K in the Tl-Ca/Ba-Cu-O system

Z. Z. Sheng & A. M. Hermann

Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, USA

The discovery of 30-K superconductivity in the La-Ba-Cu-O system¹ and 90-K superconductivity in the Y-Ba-Cu-O system² stimulated a worldwide search for even higher-temperature superconductors. Unfortunately, most of the higher-temperature transitions reported in the past year have proved to be unstable, irreproducible, or not due to bulk superconductivity³⁻⁷. Recently, we and co-workers^{8,9} reported superconductivity above 90 K in a new Tl-Ba-Cu-O system, and pointed out that elemental substitutions in this system may lead to even higher-temperature superconductivity. Here we report stable and reproducible bulk superconductivity with an onset at 120 K and zero resistance above 100 K in the Tl-Ca/Ba-Cu-O system. This transition temperature is much higher than those observed for typical rare-earth-containing superconductors, and the onset temperatures are comparable to that in the Bi-Ca/Sr-Cu-O system, as reported in refs 10 and 11 (received after submission of this paper).

A typical procedure for preparing Tl-Ca/Ba-Cu-O samples is as follows. Appropriate amounts of Ti_2O_3 , CaO and BaCu_3O_4 for a certain nominal composition (for example, $\text{Ti}_2\text{CaBaCu}_3\text{O}_{8+x}$), were completely mixed, ground and pressed into a pellet with a diameter of 7 mm and a thickness of 2-3 mm. The pellet was then put into a tube furnace, which had been heated to 880-910 °C, and was heated for 3-5 min in flowing oxygen. The sample was then taken from the furnace and quenched to room temperature in air. Some heated samples were furnace-cooled to room temperature, and some quenched samples were subsequently annealed at 450 °C in flowing oxygen for several hours. We use BaCu_3O_4 as a starting material because it has the lowest melting point among the Ba-Cu oxides, and molten BaCu_3O_4 is extremely fluid¹²; we believe that melt-solid reactions take place more completely than do typical solid-state reactions. The BaCu_3O_4 used in this experiment was prepared according to the procedure described in refs 8, 9 and 11.

Resistance was measured by the standard four-probe technique, with silver paste contacts. Rectangular bars of typical

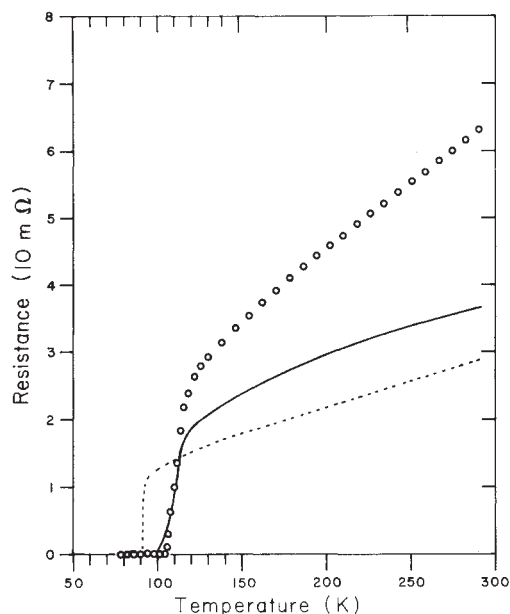


Fig. 1 Resistance versus temperature for $\text{Ti}_2\text{Ca}_{1.5}\text{BaCu}_3\text{O}_{8.5+x}$ (circles) and $\text{Ti}_{1.86}\text{CaBaCu}_3\text{O}_{7.8+x}$ (solid line), compared with $\text{EuBa}_2\text{Cu}_3\text{O}_{7-x}$ (dashed line).

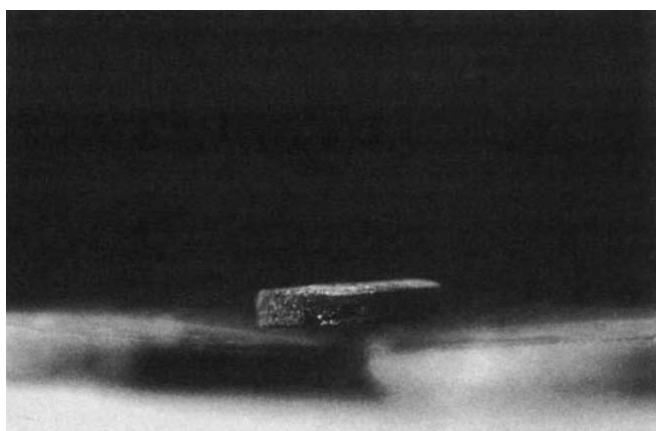


Fig. 2 A Tl-Ca/Ba-Cu-O sample levitating over a magnet after immersion in liquid nitrogen.

size $7 \times 3 \times 1$ mm were cut from the pellets for resistance measurements. The resistance of the Tl-Ca/Ba-Cu-O samples usually starts to drop sharply at ~ 120 K, and reaches zero at ~ 100 K. The midpoint of the superconducting transition is at ~ 110 K. We estimate the onset temperature as the temperature at which the curvature in the resistance versus temperature plot changes most rapidly¹³. In our data this is in approximate agreement with the onset as determined by the intersection of the slopes in the transition region and in the normal region well above the transition. Figure 1 shows resistance plotted against temperature for a furnace-cooled $\text{Ti}_2\text{Ca}_{1.5}\text{BaCu}_3\text{O}_{8.5+x}$ sample (open circles) and a $\text{Ti}_{1.86}\text{CaBaCu}_3\text{O}_{7.8+x}$ sample annealed at 450 °C (solid line). For comparison, Fig. 1 also shows the behaviour of a $\text{EuBa}_2\text{Cu}_3\text{O}_{7-x}$ sample (dashed line) which was carefully prepared using a typical solid-state sinter reaction technique. The transition midpoint and zero-resistance temperatures for the $\text{Ti}_2\text{Ca}_{1.5}\text{BaCu}_3\text{O}_{8.5+x}$ sample are 112 K and 103 K, which are the highest among the Tl-Ca/Ba-Cu-O samples prepared so far, and are much higher than those observed for high-quality rare-earth-containing samples (see, for example, refs 13 and 14), and the zero-resistance temperature is higher than those attained so far in the Bi-Sr/Ca-Cu-O system^{10,11}. The superconducting transition of the Tl-Ca/Ba-Cu-O samples is rather broad (usually > 20 K), suggesting that zero resistance could be reached at higher temperature by improvement of the preparation procedure.

Qualitative magnetic examinations of Tl-Ca/Ba-Cu-O samples showed strong diamagnetic repulsion. Figure 2 shows a sample with a nominal composition of $\text{Ti}_2\text{Ca}_{1.5}\text{BaCu}_3\text{O}_{9+x}$ levitating over a magnetic field of $\sim 5,000$ G after immersion in liquid nitrogen. Figure 3 shows the temperature dependence of the a.c. susceptibility of a sample of nominal composition $\text{Ti}_2\text{Ca}_{1.5}\text{BaCu}_3\text{O}_{8.5+x}$, as determined by the technique of Norton¹⁵. The onset temperature as determined by the deviation

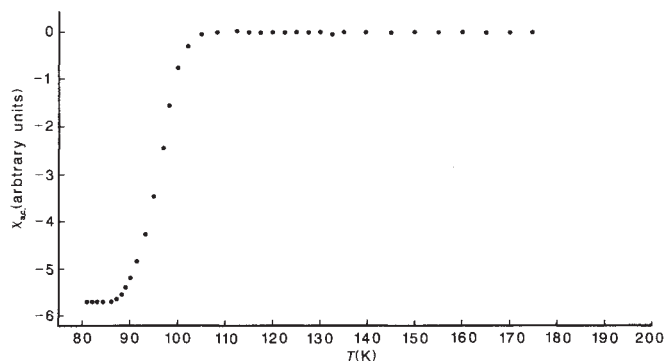


Fig. 3 Temperature dependence of the a.c. susceptibility of a $\text{Ti}_2\text{Ca}_{1.5}\text{BaCu}_3\text{O}_{8.5+x}$ sample.

from zero slope is somewhat lower than that of the initial drop in resistance—behaviour which was often seen in non-stoichiometric rare-earth superconductors (see, for example, ref. 16). Quantitative d.c. magnetization measurements are in progress.

The superconducting behaviour of the Tl-Ca/Ba-Cu-O samples depends on the preparation conditions, but not as strongly as for the Tl-Ba-Cu-O samples. Annealing at 450 °C in flowing oxygen depresses or destroys the superconducting transition of the latter whereas it slightly improves the superconductivity of the former. Variation of the Ca/Ba ratio does influence the superconducting behaviour of samples. Tl-Ca-Cu-O samples prepared under similar conditions are semiconductor-like down to liquid-nitrogen temperature, whereas Tl-Ba-Cu-O samples have zero resistance just above liquid-nitrogen temperature^{8,9}. Experiments to find an optimum Ca/Ba ratio are underway. Note that a 50% substitution of calcium for barium in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ depresses transition temperature, whereas a similar substitution in the $\text{LaBa}_2\text{Cu}_3\text{O}_{7-x}$ compound increases the superconducting transition temperature¹³.

We thank Mark J. Dreiling of the Phillips Petroleum Company for providing the low-frequency a.c. susceptibility data.

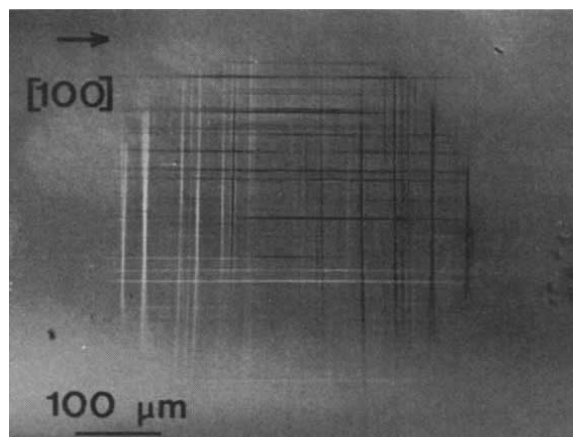


Fig. 1 Slip steps produced on the (001) surface of magnesium oxide subjected to a mean pressure of 0.7 GPa at room temperature as a result of a flattened copper cone.

Received 9 February; accepted 26 February 1988.

1. Bednorz, J. G. & Müller, K. A. *Z. Phys.* **B64**, 189–193 (1986).
2. Wu, M. K. *et al. Phys. Rev. Lett.* **58**, 908–910 (1987).
3. Chen, J. T., Wenger, L. E., McEwan, C. J. & Logothetis, E. M. *Phys. Rev. Lett.* **58**, 1972–1975 (1987).
4. Ovshinsky, S. R., Young, R. T., Allred, D. D., DeMaggio, G. & Van der Leeden, G. A. *Phys. Rev. Lett.* **58**, 2579–2581 (1987).
5. Cai, X., Joynt, R. & Larbalestier, D. G. *Phys. Rev. Lett.* **58**, 2798–2801 (1987).
6. Bhargava, R. N., Herko, S. P. & Osborne, N. *Phys. Rev. Lett.* **59**, 1468–1471 (1987).
7. Huang, C. Y. *et al. Nature* **328**, 403–404 (1987).
8. Sheng, Z. Z. & Hermann, A. M. *Nature* **332**, 55–58 (1988).
9. Sheng, Z. Z. *et al. Phys. Rev. Lett.* (in the press).
10. Maeda, H., Tanaka, Y., Fukutomi, M. & Asano, T. *Jap. J. appl. Phys. Lett.* (submitted).
11. Chu, C. W. *et al. Phys. Rev. Lett.* (submitted).
12. Hermann, A. M. & Sheng, Z. Z. *Appl. Phys. Lett.* **51**, 1854–1856 (1987).
13. Murphy, D. *et al. Phys. Rev. Lett.* **58**, 1888–1890 (1987).
14. Hor, P. H. *et al. Phys. Rev. Lett.* **58**, 1891–1894 (1987).
15. Norton, N. L. *J. Phys. E* **19**, 268–270 (1986).
16. Dietrich, M. R., Fietz, W. H., Ecke, J., Odest, B. & Politis, C. *Z. Phys.* **B66**, 284–287 (1987).

Multiple slip in diamond due to a nominal contact pressure of 10 GPa at 1,000 °C

C. A. Brookes, V. R. Howes & A. R. Parry

Department of Engineering Design and Manufacture, University of Hull, Hull HU6 7RX, UK

From studies of anisotropy in Knoop hardness¹ and transmission electron microscopy of the deformation beneath indentations², it is widely accepted that some dislocation movement is possible in diamond at room temperature. Three-point bend tests were used earlier to identify a brittle-to-ductile transition temperature of 1,500 °C (ref. 3), and extensive dislocation movement was considered unlikely at lower temperatures. Here we show, however, that multiple slip can occur in diamond as a result of the pressure transmitted by the point-loading of a softer material, namely a cubic boron nitride cone at 1,000 °C. The method enables specimens to be prepared for electron microscopy and X-ray topography that have a controlled density of dislocations introduced at homologous temperatures below $0.3T_m$, where T_m is the melting point. Furthermore, it underlines the importance of microplasticity in the application and performance of crystalline engineering ceramics.

In the deformation of hard ceramic crystals at low homologous temperatures, catastrophic brittle fracture often dominates

because dislocation mobility is usually very limited. Although the mechanisms of slip and brittle fracture are competitive, they are not mutually exclusive. Thus, restricted dislocation movement can cause fracture on {110} planes, rather than on the normal {100} cleavage planes in magnesium oxide⁴. Also, the suggestion that limited dislocation movement can control the mechanical behaviour of these crystals is supported by the high degree of correlation between their markedly anisotropic Knoop hardness and models based on resolved shear stresses on active slip systems^{5,6}.

Encouraging plastic deformation, rather than fracture, in the covalent crystals is more difficult than in most ionic, and in all metal, crystals for a number of reasons. First, the nature of the crystal bonding and its pronounced directionality leads to high Peierls' stresses which resist dislocation motion. Thus, their critical resolved shear stresses are likely to be in the region of $10^{-2}G$ (where G is the bulk shear modulus), rather than the $10^{-5}G$ typical of metals. Second, the critical resolved shear stress for a given covalent crystal is much more dependent on the strain rate than it is in other solids—for example, in silicon it is increased by a factor of 2 for a change in strain rate of a factor of 10. Third, and directly related to the comparative difficulty of dislocation movement, they have low-fracture toughnesses. Nevertheless, there is generally a critical temperature above which dislocation mobility is significantly increased, and this temperature has been determined using indentation techniques in conjunction with dislocation etch pitting and X-ray topography⁷. By both methods it was shown that the critical temperature for enhanced dislocation expansion beneath Vickers indentations was ~ 350 °C for undoped germanium crystals. In contrast, dislocation etch pit patterns around indentations in germanium have recently been found to enlarge more gradually with temperature, but the extent of radial cracks on the surface decreased sharply at the same temperature⁸. Similarly, the threshold temperature necessary for the spontaneous expansion of dislocation rosettes around room-temperature indentations that have been subsequently annealed has been identified for germanium (300 °C), silicon (400 °C) and cubic boron nitride (900 °C)⁹. It has been suggested that the threshold temperature for diamond is below 1,100 °C, on the basis of slip step formation around indentations⁹, but the brittle-to-ductile transition temperature of 1,500 °C from three-point bend tests³ has generally been recognized as the minimum temperature for multiple slip in diamond¹⁰.

A novel method of encouraging dislocation movement, rather than brittle fracture, has been exploited recently to demonstrate that hard crystals like magnesium oxide and titanium carbide

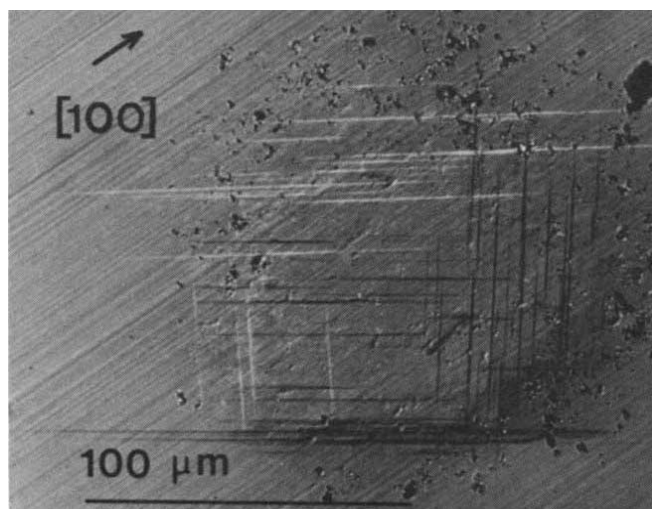


Fig. 2 Slip steps produced on the near (001) surface of a Type I diamond subjected to a mean pressure of 10 GPa at 1,000 °C as a result of a flattened cubic boron nitride cone. Note the diagonal lines across the micrograph aligned in a [100] direction are a feature of the polishing process; there is some debris from the cone adherent to the contact area.

can undergo localized plastic deformation and work-hardening when repeatedly traversed by softer lubricated metal sliders at room temperature¹¹. To induce this behaviour, the nominal contact stress should exceed the critical resolved shear stress but not the fracture stress of the crystal. In practice, the tip of the softer slider macroscopically plastically deforms until its contact area is sufficient to support the applied load. (The applied load divided by the nominal contact area gives a mean pressure, P_m , which is typically about one third of the measured hardness of the plastically deformed tip for metals). Thus, for a particular orientation of the slip systems and cleavage planes with respect to the deformed surface, control over the magnitude of the applied stress can be exercised through the properties of the material used for the slider. If the mean pressure developed by a given slider material is only just sufficient to exceed the critical resolved shear stress within the bulk of the harder crystal, then relatively few dislocations will be moved during the first traversal. Further traversals will induce additional dislocation movement and the consequent work-hardening quickly increases the flow stress of the crystal to a level which enables the mean pressure to be supported elastically. The level of this work-hardening is then directly related to the hardness of the slider, thereby inducing a specific strain in the crystal, but it can be curtailed ultimately by fatigue failure.

The use of the known properties of one material to evaluate those of another is particularly relevant to the hardest materials. We have therefore studied the nature of deformation when a softer cone (initially formed with a 120° included angle) is used to apply a normal load to the surface of a hard crystal. The apparatus used¹² enables a normal load to be applied gradually from zero to 115 N, this load representing the difference in pressure between atmospheric and the vacuum of 10^{-5} mbar maintained in the specimen chamber. Hence, the maximum pressure that can be applied to the hard crystal cannot exceed the flow stress of the material used for the softer cone. This material is chosen so that it is incapable of transmitting a pressure that causes cracking rather than slip. Higher loads can be applied by additional weights attached directly to the indenter support column. The initial experiments on magnesium oxide crystals take advantage of the wide basic knowledge available for interpreting the deformation of this particular material. First a polycrystalline copper cone was used to deform a (001) mag-

nesium oxide surface. On the simplest assumptions, this system represents a compressive load normal to the surface which, ignoring edge effects and the difference in the elastic modulus of the two materials, can be resolved on the {110}<110> slip systems using a Schmid-Boas factor of 0.5. Alternatively, a recent model (S. G. Roberts, personal communication)¹³, based on the resolution of shear stresses beneath a circular contact with uniform stress, predicts a maximum shear stress at a position beneath the surface at a depth of ~ 0.33 of the contact diameter (a) and 0.35 of the nominal pressure (P_m). Hence, given that the magnesium oxide has a critical resolved shear stress of 0.06 GPa (ref. 11), the measured mean pressure due to the blunted copper tip ($P_m = 0.7$ GPa) develops a resolved shear stress (0.25–0.35 GPa) well in excess of that critical level. Consequently, slip lines are formed on the MgO surface, corresponding to the emergence of edge dislocations on (011) and (101) slip planes, in the contact area (Fig. 1). When a phosphor-bronze cone is used, the mean pressure ($P_m = 1.4$ GPa) is sufficient to produce slip lines both in the area of contact and on the magnesium oxide surface well outside that region. At a still higher level of mean pressure (a stainless steel cone, $P_m = 4$ GPa) there is not only a greater amount of slip in both of these regions but, immediately adjacent to the contact area, cracks are also produced on those {110} planes normal to the surface. These are the cracks referred to earlier, thought to be due to dislocation interactions on intersecting slip planes⁴. Subsequently, the same technique has been used to investigate the possibility of inducing multiple slip in diamond at a temperature well below the normally accepted brittle-ductile transition, but at which indentation creep has previously been reported, at 1,000 °C (ref. 13). In these experiments, a cone was made from an aggregate of cubic boron nitride and the experiments were carried out in a vacuum of $\sim 10^{-5}$ mbar using a polished Type I diamond surface which was a few degrees off the (001) plane. The cone flattened after a loading time of ~ 5 min, such that the applied pressure was ~ 10 GPa. This pressure was sufficient to develop a considerable number of slip steps within the contact area (Fig. 2), directly comparable with those for magnesium oxide and consistent with multiple slip on intersecting {111}<110> slip systems. There were no cracks formed under these particular conditions. The model referred to earlier (S. G. Roberts, personal communication) predicts that the maximum shear stress is $\sim 0.32 P_m$, for a (001) surface with {111}<110> slip systems, and lies on a circular line with a radius of $0.27a$; this is concentric with the centre line of the contact area and $\sim 0.29a$ below the surface. Clearly, more meaningful estimates of the critical resolved shear stresses must await a better understanding of the effective strain rates developed under these particular experimental conditions. Nevertheless, multiple slip can probably be induced in diamond at 1,000 °C under a resolved shear stress of ~ 3.2 GPa. It was also noted that under the same experimental conditions, but at higher loading pressures (~ 15 GPa), some cracking was produced outside the contact area in the same way as for magnesium oxide. There is a dearth of information on the critical resolved shear stress for plastic flow in diamond because of the associated experimental difficulties. Indeed, the only data are those obtained during three-point bending of Type I and Type II diamonds at temperatures above 1,500 °C, below which the diamonds fractured rather than plastically deformed³, when calculated values of the critical resolved shear stress at 1,800 °C ranged from 0.5 GPa to 1.2 GPa for Type II and Type I diamonds respectively. In comparing those values with the resolved shear stress estimated here for multiple slip, we should bear in mind not only the possible effect of strain rate, but also the fact that these experiments were carried out at homologous temperatures above $0.5T_m$, where presumably thermal recovery would be significant.

This demonstration of the way conditions of stress can be manipulated to encourage plastic deformation, rather than brittle fracture, extends to all crystalline solids, but is likely to

be particularly useful for studying deformation in diamond and other hard ceramic crystals. In addition to affording a simple method for estimating the resolved shear stresses required to cause plastic flow in covalent solids that are difficult not only to deform in uniaxial stress condition, but also to shape, it permits control over the density and distribution of mobile dislocations introduced at low homologous temperatures. Consequently, techniques such as electron microscopy and X-ray topography can now be applied to study dislocation interactions and work-hardening of these materials for direct comparison with ionic and metallic solids. Application of these principles could also be used to modify both the topography and the integrity of the surfaces of components made from engineering ceramics, such as the healing of surface defects by localized plastic flow.

We thank De Beers Industrial Diamond Division Ltd for support, and the Royal Society for a Visiting Research Fellowship for one of us (V.R.H.) on sabbatical from the University of New South Wales. We also thank Professor Sir Peter Hirsch for his comments on the original manuscript.

Received 15 December 1987; accepted 5 February 1988.

1. Brookes, C. A. *Nature* **228**, 660–661 (1970).
2. Humble, P. & Hannink, R. H. *Nature* **273**, 37–39 (1978).
3. Evans, T. & Wild, R. K. *Phil. Mag.* **12**, 479–486 (1965).
4. Keh, A. S., Li, J. C. M. & Chou, Y. T. *Acta Metall.* **7**, 694–699 (1959).
5. Brookes, C. A., O'Neill, J. B. & Redfern, B. A. W. *Proc. R. Soc. A* **322**, 73–88 (1971).
6. Roberts, S. G., Warren, P. D. & Hirsch, P. B. *J. Mater. Sci.* (in the press).
7. Kent, E. J., University Bristol (1976).
8. Hirsch, P. B., Pirousz, P., Roberts, S. G. & Warren, P. D. *Science of Hard Materials* (eds Almond, E. A., Brookes, C. A. & Warren, R.) 83–105 (Inst. Phys. Conf. Ser. No. 75, Hilger, Bristol, 1988).
9. Brookes, C. A., Hooper, R. M., Lambert, W. A. & Ross, J. D. J. *Ultrahard Materials Applications Technology* Vol. 2 (ed. Daniel, P. A.) 72–80 (De Beers International Diamond Div., 1983).
10. Frank, F. G., Lawn, B. R., Lang, A. R. & Wilks, E. M. *Proc. R. Soc. A* **301**, 239–252 (1967).
11. Brookes, C. A., Shaw, M. P. & Tanner, P. E. *Proc. R. Soc. A* **409**, 141–159 (1987).
12. Harrison, P. thesis, Univ. Exeter (1973).
13. Brookes, C. A. *Properties of Diamond* (ed. Field, J. E.) 383–402 (Academic, London, 1979).

Enclathration of helium in ice II: the first helium hydrate

D. Londono^{*†}, W. F. Kuhs^{*‡} & J. L. Finney[†]

^{*} Institut Laue–Langevin, BP 156X, I-38042 Grenoble Cedex, France

[†] Birkbeck College, Malet Street, London WC1E 7HX, UK

The hydrates of the heavier noble gases Ar, Kr and Xe are easy to prepare, and they crystallize in two different structures¹. The corresponding hydrates of He and Ne apparently do not exist. We now find by neutron powder diffraction experiments that, when applying He gas pressures of >0.28 GPa, a gas hydrate with the idealized formula He · 6D₂O is formed. The He content varies with the applied He pressure as expected of an ideal solution behaviour with a Langmuir constant of 0.07 (1) GPa⁻¹. The host lattice of this first well-characterized He-clathrate is very similar to ice II. The He enclathration stabilizes the ice II structural framework and hampers the formation of ice III, V and IX.

Gas hydrates of the noble gases were first described by Villard² (argon) and de Forcrand^{3,4} (krypton and xenon). Most of the gas hydrates crystallize in one of two distinct clathrate structures: type I provides cages for molecules up to 5.8 Å in diameter while type II accepts molecules with a maximum dimension of 7.0 Å. Both types of structure also have smaller cages of ~5.0 Å diameter. For a long time it was assumed that all three noble-gas hydrates crystallize in the type I structure and only recently has it been realized that argon and krypton enter into the small

cages of type II clathrate¹. The lighter noble gases (obviously) do not form clathrates of type I or II. Only for neon has it been shown⁵ that it can enter in small quantities into a clathrate type II structure where it acts as a stabilizing hilfs-gas (help-gas). The main reason for this size-dependent behaviour is the relative instability of the host-lattice when the cavities are either not occupied or are occupied by atoms or molecules that are too small, and which, in addition, escape easily from the lattice⁶. Consequently, the corresponding pure neon or helium clathrates would be completely unstable under near-ambient conditions.

On the other hand, helium and neon are found to be sorbed in tridymite and cristobalite⁷. These silicate structures are topologically identical to those of ice Ih and ice Ic respectively and possess rather open channels, which presumably accept small gas species with van der Waals radii smaller than ~1.5 Å. Likewise it might be suspected that such small atoms could enter into the lattice of ice Ih and indeed there is experimental evidence that this happens to a very limited extent^{8,9}. In a series of neutron powder diffraction experiments on heavy ices under helium-pressures between ambient and 0.5 GPa we have now found positive evidence for the formation of a non-stoichiometric helium gas hydrate with the ideal composition He · 6D₂O.

Powders of ice Ih were prepared by grinding frozen heavy water (99.5% purity) in a mortar at liquid-nitrogen temperatures. The samples were filled in vanadium cans and mounted into a helium pressure cell¹⁰, which itself was inserted in a variable temperature cryostat. The loaded cryostat was placed on to the D1A high-resolution powder diffractometer at the Institut Laue–Langevin, Grenoble. Thus, structural changes could be monitored as a function of pressure and temperature. A phase transformation from ice Ih into the helium hydrate was observed at 0.28 (1) GPa and 252 (2) K. The powder pattern of this new phase resembles closely the pattern of ice II and indeed the hydrogen-bonded networks of the water molecules in the two phases are closely similar. Two complete data sets were taken at 0.290 (5) and 0.470 (5) GPa and 195 (1) K. Structural parameters were obtained from Rietveld-profile refinements.

As in ice II (ref. 11) the hydrogen (deuterium) atoms are fully ordered. Intra- and intermolecular distances are given in Table 1. There are two types of water molecules in the structure, each type building up 6-membered rings (Fig. 1). The ring formed by water-2 molecules is rather flat, while the other ring formed by water-1 molecules is more puckered; this latter ring connects to the flatter water-2 ring by donating protons in hydrogen bonds

Table 1 Selected intra-, inter-molecular and interatomic distances and angles in He · (6+x)D₂O at 0.29 GPa and 195 K (space group: R $\bar{3}$; lattice constants: $a_0 = 12.934(1)$, $c_0 = 6.216(1)$ Å)

Intra-molecular					
O1—D1	0.980 (12)	D1—O1—D2	102.5 (1.0)		
O1—D2	0.985 (8)				
O2—D3	0.973 (10)	D3—O2—D4	105.5 (0.8)		
O2—D4	0.958 (7)				
Inter-molecular					
O1...O1	2.769 (18)	O1—D1...O1	166.5 (0.8)		
O1...O2	2.815 (9)	O1—D2...O2	167.3 (0.6)		
O2...O2	2.765 (13)	O2—D3...O2	173.6 (0.7)		
O2...O1	2.801 (8)	O2—D4...O1	172.4 (0.9)		
Non-bonded					
O1...O2	3.222 (8)				
O2...O2	3.533 (9)				
Helium environment					
He...O1	3.060 (32)	He...D1	2.933 (40)	He...D2	2.876 (13)
	3.370 (41)		2.872 (39)		4.098 (49)
He...O2	3.043 (29)	He...D3	2.745 (32)	He...D4	3.985 (31)
	3.137 (33)		2.846 (36)		3.705 (19)
He...He	2.783 (94)				
	3.432 (94)				

[‡] Author to whom correspondence should be sent at present address: Institute of Crystallography, University of Karlsruhe, Postfach 6380, D-7500 Karlsruhe 1, FRG.

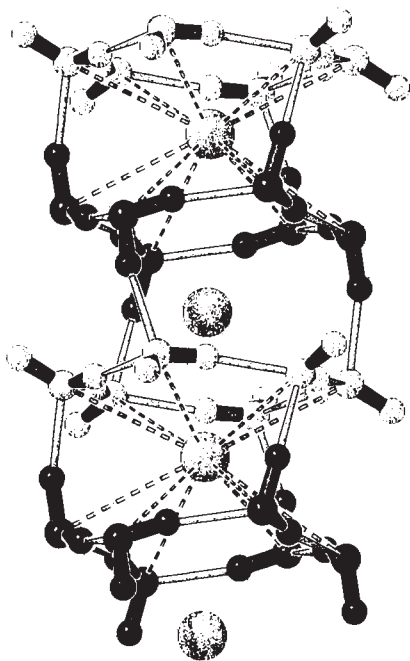


Fig. 1 Detail of the structure of $\text{He} \cdot (6+x)\text{D}_2\text{O}$ at 0.47 GPa and 195 K showing the helium atoms occupying sites approximately halfway between the two types of 6-membered rings (in black the water-1, in white the water-2 molecules). The channel runs along the rhombohedral axis (drawn by SCHAKAL written by E. Keller, University of Freiburg, FRG).

along the rhombohedral axis. The resulting structure with channels running along the unique axis provides voids of sufficient size to host small atoms or molecules. The helium atoms are found in positions approximately halfway between the 6-membered rings with a slight shift towards the flatter ring formed by water-2 molecules. Thus there is one shorter and one longer interatomic helium-helium distance. The occupancy ϕ_{He} of the helium positions was found to increase as the pressure increases (61.8% occupancy at 0.29 GPa, 78.8% occupancy at 0.47 GPa). Such an increase is in agreement with an ideal solution behaviour as discussed by van der Waals and Platteeuw¹² where

$$\phi_{\text{He}} = C_{\text{He}} \cdot P_{\text{He}} / (1 + C_{\text{He}} \cdot P_{\text{He}})$$

with P_{He} being the helium fugacity. The Langmuir constant of helium C_{He} which fits the observations is $0.07 \text{ (1 GPa)}^{-1}$.

From our experiments, we can draw the following conclusions concerning the phase stability of the helium hydrate. (1) Once formed it apparently cannot be transformed into any phase other than the liquid (on warming) or ice Ih (on pressure reduction) at least up to 0.48 GPa. Thus it occupies the region of ices III, V and IX in the corresponding helium-free ice phase diagram. (2) Helium hydrate is formed from the liquid on cooling under helium pressure (for example the melting point $T_{\text{M}} = 254 \text{ (2 K)}$ at 0.480 (5) GPa). (3) Remarkable also is the high transformation pressure from ice Ih of $\sim 0.28 \text{ GPa}$, a pressure about 0.06 GPa higher than that at which helium-free ice Ih transforms to ice II according to the generally accepted phase diagram for pure ice^{13,14}.

Qualitatively, we can begin to understand these phase boundary shifts in terms of the structures of the related ice phases. It has already been suspected that helium enters into the ice Ih lattice¹⁵ thus changing the phase boundaries between the different phases and the work reported here provides evidence to confirm this. Ices III (IX) and V do not have the wide channels found in ice Ih and ice II and it is therefore conceivable that once helium has entered the ice Ih lattice the consequent necessity of expelling the helium gas before transforming to ices III (IX) and V will suppress these transformations. Apparently,

the difficulty of expelling the helium from ice Ih means that when it eventually transforms, it does so into the helium hydrate with an ice II-like host lattice conserving essentially its helium content. We have indirect evidence to support this suggestion in similar experiments in which the helium is excluded from ice contact by encapsulation in indium containers. There, we observe normal phase transformation behaviour, which is in agreement with other experiments we have undertaken in a clamped pressure cell with helium completely absent. Under both these sets of conditions ices II and III are formed at pressures and temperatures^{13,14} that are consistent with the generally accepted ice phase diagram.

We thank the ILL for neutron facilities and the SERC and the Birkbeck College Research Committee for financial assistance. We are grateful to Dr J. Cockcroft (MPI Stuttgart) for providing the programs to perform the Rietveld profile refinements.

Received 10 November 1987; accepted 5 January 1988.

1. Davidson, D. W., Handa, Y. P., Ratcliffe, C. I., Tse, J. S. & Powell, B. M. *Nature* **311**, 142-143 (1984).
2. Villard, P. C. *r. heb. Séanc. Acad. Sci., Paris* **123**, 377-379 (1896).
3. de Forcrand, R. C. *r. heb. Séanc. Acad. Sci., Paris* **176**, 355-358 (1923).
4. de Forcrand, R. C. *r. heb. Séanc. Acad. Sci., Paris* **181**, 15-17 (1925).
5. Barrer, R. M. & Ruzicka, D. J. *Trans. Faraday Soc.* **58**, 2239-2252 (1962).
6. Child, W. C. *Quart. Rev. Chem. Soc.* **18**, 321-346 (1964).
7. Barrer, R. M. & Vaughan, D. E. W. *Trans. Faraday Soc.* **63**, 2275-2290 (1967).
8. Kahane, A., Klinger, J. & Philippe, M. *Solid St. Commun.* **7**, 1055-1056 (1969).
9. Namiot, A. Y. & Bukhgalter, J. *Struct. Chem.* **6**, 873-874 (1965).
10. Poireau, J. & Vettier, C. *Rev. sci. Instrum.* **46**, 1484-1488 (1975).
11. Kamb, B., Hamilton, W. C., La Placa, S. J. & Prakash, A. *J. chem. Phys.* **55**, 1934-1945 (1971).
12. Van der Waals, J. H. & Platteeuw, J. C. *Adv. chem. Phys.* **2**, 1-5 (1959).
13. Tammann, G. *Kristallisieren und Schmelzen* 315-344 (Barth, Leipzig, 1903).
14. Bridgman, P. W. *J. chem. Phys.* **3**, 597-605 (1935).
15. Arnold, G. P., Wenzel, R. G., Rabideau, S. W., Nereson, N. G. & Bowman, A. J. *chem. Phys.* **55**, 589-595 (1971).

Lidar observations of substantial sodium depletion in the summertime Arctic mesosphere

C. S. Gardner, D. C. Senft & K. H. Kwon

Department of Electrical and Computer Engineering,
University of Illinois at Urbana-Champaign,
1406 West Green Street EEB, Urbana, Illinois 61801, USA

The chemistry and dynamics of the mesospheric sodium layer have been studied extensively with lidar (laser radar) techniques since the late 1960s. Meteoric ablation is generally regarded as the dominant source of all mesospheric alkali metals including sodium. The Na layer is typically confined to the region between 80 and 110 km with a peak near the mesopause at 90 km, where the density ranges from $\sim 10^3$ to 10^4 cm^{-3} . The Na column abundance at mid-latitudes in the Northern Hemisphere varies from a summer minimum of $\sim 3 \times 10^9 \text{ cm}^{-2}$ to a winter maximum of $\sim 10^{10} \text{ cm}^{-2}$ in December and January¹. The seasonal and geographical variations in Na abundance are now believed to be related to changes in the mesopause temperature that affect the reaction rates of the main chemical loss processes for Na. In addition to chemical activity, the dynamic effects of the tides, gravity waves and the mean winds have a significant influence on the vertical structure of the layer. In June, the University of Illinois (UIUC) sodium lidar system was installed at Nordlysstasjonen, Svalbard ($78^\circ 12' \text{ N}$, $15^\circ 50' \text{ E}$) on Spitsbergen Island and we conducted more than 27 h of observations during the period 10-15 July 1987. The average Na layer column abundance was $\sim 6 \times 10^8 \text{ cm}^{-2}$, a value that is almost five times lower than typical summertime Na abundances measured at mid-latitudes. Additional observations were conducted during the period 7-11 September 1987 and the measured Na abundances varied between $\sim 5 \times 10^9$ and $14 \times 10^9 \text{ cm}^{-2}$.

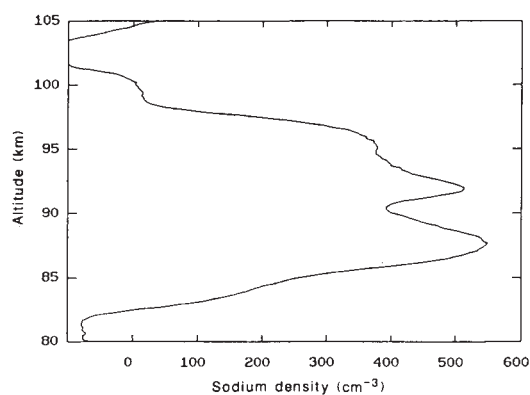


Fig. 1 Average sodium density profile above Nordlysstasjonen, Svalbard during the period 10–15 July 1987. The data were smoothed using modified Hamming window with full-width-half-maximum = 6 km.

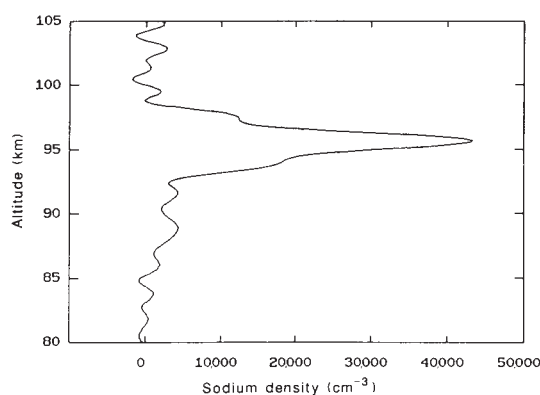


Fig. 2 Sodium density profile above Nordlysstasjonen, Svalbard at 21:52 UT 7 September 1987. The integration period was 14 min and the data were spatially smoothed using a low-pass filter with a cutoff frequency of 0.6 km^{-1} .

The system parameters and daytime capabilities of the UIUC Na lidar are described by Kwon *et al.*². Data processing and analysis are discussed by Gardner *et al.*¹. Figure 1 is a plot of the average Na profile measured during the period 10–15 July 1987 at Nordlysstasjonen, Svalbard. Figures 2 and 3 are plots of the profiles measured near midnight local time (LT) on 7 and 9 September 1987 respectively. Table 1 lists the main sodium layer parameters for these three profiles. The large standard deviations in the parameters measured during July are caused by the low sodium signal level and the high background noise from the bright sunlit sky.

Although in July the centroid height and r.m.s. thickness are comparable with values measured at other locations, the peak density and column abundance are almost a factor of five lower. These results, which to our knowledge are the first sodium measurements made near the North Pole during summer, are very surprising, particularly when compared with other high-

ing the September campaign. The column abundance ($\sim 4.8 \times 10^9 \text{ cm}^{-2}$) is comparable with mid-latitude values that have been measured in the Northern Hemisphere in early September¹.

The chemistry of the Na layer is still poorly understood even though numerous models have been proposed in an attempt to explain the general features of the seasonal, diurnal and geographical variations in the layer. Models for the dominant loss mechanisms have generally evolved in two directions. One group of models has been developed by assuming neutral reactions dominate Na chemistry^{8,9}. The other group assumes that ionic species, particularly cluster ions of the form $\text{Na}^+(\text{H}_2\text{O})_n$, are of considerable importance^{10,11}. A complete understanding of the Na layer chemistry has been hampered by the lack of reliable estimates of the rate coefficients for many of the important reactions. Even so, it is generally agreed that the seasonal variation in mesopause temperature is primarily responsible for the seasonal variation of Na abundance at mid-latitudes.

Swider¹² recently reported model calculations which suggest that the neutral reaction $\text{Na} + \text{O}_2 + \text{M} \rightarrow \text{NaO}_2 + \text{M}$, where M is any molecule, is an important sink for mesospheric Na. The Na loss rate due to this reaction has a T^{-4} temperature dependence. Von Zahn and Neuber¹³ report that the winter mean temperature between 80 and 100 km at Andøya (69°N) ranges from ~ 200 to 220 K. In early August 1982, Philbrick *et al.*¹⁴ measured temperatures for the same altitude region at Kiruna, Sweden (68°N). Their values were in the range from ~ 100 to 140 K. Thus the Na loss rate at high latitudes for the neutral reaction discussed by Swider could be as much as ten times larger in summer compared to winter. However, because the July sodium abundance at Nordlysstasjonen is almost 24 times smaller than the 7 September levels and almost 9 times smaller than the 9

Table 1 Sodium layer parameters measured at Nordlysstasjonen, Svalbard ($78^\circ 12' \text{N}$, $15^\circ 15' \text{E}$)

Observation period	21:59 UT 10 July–01:00 UT 15 July 1987	21:52–22:06 UT 7 Sept. 1987	00:00–00:40 UT 9 Sept. 1987
Peak density (cm^{-3})	550 ± 50	$(4.3 \pm 0.2) \times 10^4$	$(7.8 \pm 0.4) \times 10^3$
Column abundance (cm^{-2})	$(5.7 \pm 2) \times 10^8$	$(1.4 \pm 0.1) \times 10^{10}$	$(4.8 \pm 0.1) \times 10^9$
Centroid height (km)	90.6 ± 2	94.6 ± 0.2	93.5 ± 0.1
r.m.s. thickness (km)	3.82 ± 2	2.52 ± 0.3	3.23 ± 0.1

latitude observations. It is well known that Na abundance is quite high in the wintertime polar mesosphere. Lidar measurements by Megie *et al.*³ and Juramy *et al.*⁴ at Heyss Island, Franz Joseph Land (80°N , 50°E), by von Zahn and Tilgner⁵ at Andøya, Norway (69°N , 16°E) and by Nomura *et al.*⁶ at Syowa Station, Antarctica (69°S , 40°E) show Na abundances ranging from 3×10^9 to $12 \times 10^9 \text{ cm}^{-2}$ during the polar winter. Von Zahn *et al.*⁷ report that Na abundance in July at Andøya averaged approximately $6 \times 10^8 \text{ cm}^{-2}$, which is consistent with our July measurements at Nordlysstasjonen $\sim 1,000 \text{ km}$ to the north.

The 7 September profile plotted in Fig. 2 is also interesting because of the high peak density and small thickness of the layer. The column abundance for this profile is almost 24 times larger than the abundance measured in July. The 2-km half-width and high peak density are characteristic of the very narrow layers which von Zahn *et al.*⁷ have observed on occasion forming rapidly above Andøya. The profile plotted in Fig. 3, is more typical of the layer structure observed at Nordlysstasjonen dur-

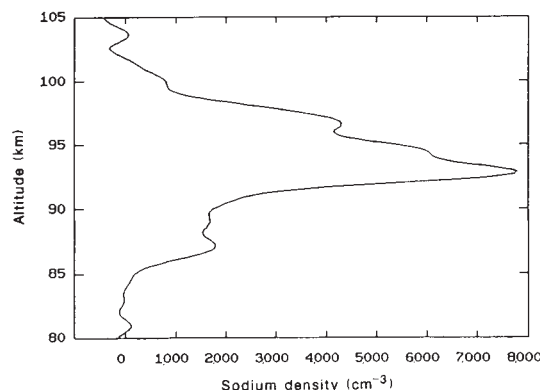


Fig. 3 Sodium density profile above Nordlysstasjonen, Svalbard at midnight 9 September 1987. The integration period was 40 min and the data were spatially smoothed using a low-pass filter with a cutoff frequency of 0.6 km^{-1} .

September levels, this process does not appear to be completely responsible for the observed summertime depletion above Svalbard. Other loss mechanisms which may play significant roles in Na depletion at high latitudes include absorption by polar mesospheric cloud particles, absorption on smoke or dust particles and photoionization.

Although the very cold temperatures near the summertime Arctic mesopause will increase the effectiveness of the dominant chemical loss processes for Na, this mechanism does not appear to be strong enough to explain the large depletion we observed. Additional measurement campaigns at Nordlysstasjonen are planned for the late autumn and winter 1987–88.

We acknowledge the valuable logistical support provided during the July and September campaigns by Kjell Henriksen of the University of Tromsø, Leif Eldring, Governor of Svalbard and Cesar LaHoz of the European Incoherent Scatter Scientific Association. This work was supported by the NSF.

Received 24 August; accepted 29 December 1987.

1. Gardner, C. S., Voelz, D. G., Sechrist, C. F. Jr & Segal, A. C. *J. geophys. Res.* **91**, 13659–13673 (1986).
2. Kwon, K. H., Gardner, C. S., Senft, D. C., Roesler, F. L. & Harlander, J. *J. geophys. Res.* **92**, 8781–8786 (1987).
3. Megie, G., Chanin, M. L., Tulinov, G. Y. & Doudoladov, Y. P. *Planet. Space Sci.* **26**, 509 (1978).
4. Juramy, P., Chanin, M. L., Megie, G., Tulinov, G. F. & Doudoladov, Y. P. *J. atmos. Terr. Phys.* **43**, 209–215 (1981).
5. von Zahn, U. & Tilgner, C. *Proc. 8th Eur. Space As. Symp. Eur. Rocket and Balloon Prog. ESA SP-270* Sunne, Sweden, 17–23, May 1987.
6. Nomura, A. & Kano, T. *Geophys. Res. Lett.* **14**, 700–703 (1987).
7. von Zahn, U., Hansen, G. & Kurzawa, H. *Nature* **331**, 594–596 (1988).
8. Kirchhoff, V. W. J. H. *Can. J. Phys.* **64**, 1664 (1986).
9. Thomas, L., Isherwood, M. C. & Bowman, M. R. *J. atmos. Terr. Phys.* **45**, 587–593 (1983).
10. Richter, E. S. & Sechrist, C. F. Jr. *J. atmos. Terr. Phys.* **41**, 579–586 (1979).
11. Jegou, J. P., Granier, C., Chanin, M. L. & Megie, G. *Ann. Geophys.* **3**, 299–312 (1985).
12. Swider, W. *Geophys. Res. Lett.* **12**, 589–592 (1985).
13. von Zahn, U. & Neuber, R. *Beitr. Phys. Atmos.* **60**, 294 (1987).
14. Philbrick, C. R. *et al. Adv. Space Res.* **4**, 153–156 (1984).

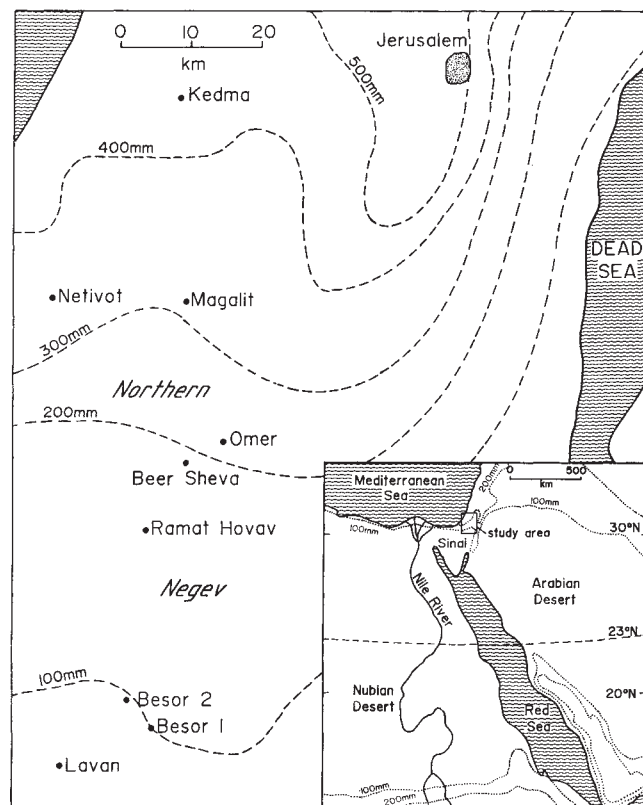


Fig. 1 Map showing location of the study sites and the location of the study area in relation to the Saharan-Arabian desert belt (inset). Present mean annual rainfall isohyets are shown. Not shown is the site of Talme El'azar, located in the coastal plain north of the area shown in the map.

Palaeosols and late Pleistocene rainfall fluctuations in the Negev Desert

Glenn A. Goodfriend & Mordechai Magaritz

Isotope Department, Weizmann Institute of Science, 76100 Rehovot, Israel

The existence of a series of palaeosols in the present-day desert region of southern Israel (the northern Negev Desert)^{1,2} indicates recurring periods of wetter climate in the past. Up to seven well-developed palaeosols occur within a section and each has a horizon containing soft calcium carbonate nodules (3–15 mm in diameter). The palaeosols are developed within loess or fluviially redeposited loess consisting of silty clay or clayey silt. The ¹⁴C ages of the uppermost three calcic horizons average ~13,000 yr BP (calcic horizon, CH-I), ~28,000 yr BP (CH-II), and ~37,000 yr BP (CH-III). Drier phases occurred between these periods of pedogenesis, especially during the last glacial maximum (centred at ~18,000 yr BP), when extensive erosion occurred. Analysis of the ¹³C content of nodule carbonates points to the occurrence of strong north-south rainfall gradients during each period of calcic horizon formation, which excludes the possibility of monsoonal rains contributing to the increased rainfall during these periods. The palaeosols correspond to global warm phases of the last glacial period.

Dating and carbon isotope analysis (¹³C/¹²C of carbonates) of the uppermost three palaeosols was carried out at nine sites in the Northern Negev (Fig. 1) and one site on the north coastal plain of Israel. The sections (2–7 m in height) were exposed by road cuts, wadi cuts, or by excavations. The time of formation of the soil nodules was determined by ¹⁴C analysis of nodule carbonate. The formation of nodules involves dissolution of

carbonate in the soil by water charged with CO₂ from the soil atmosphere and precipitation of carbonate lower down in the soil as the water evaporates or loses CO₂. The dates (Table 1) show stratigraphic consistency within each site where multiple horizons were analysed for ¹⁴C. A range of radiocarbon ages was obtained for each of these calcic horizons: 11,000–15,000 yr BP for CH-I, 23,000–31,000 yr BP for CH-II, and 34,000–39,000 yr BP for CH-III. The date of 19,500 yr BP at the Besor I site is tentatively rejected owing to inconsistency with other CH-I dates; it may have contained some limestone fragments. Radiocarbon dating of soil carbonate nodules is subject to several possible errors: the apparent ages may be too great because (1) the nodule may contain some ancient detrital carbonate, or (2) the pedogenic carbonate itself may be precipitated from soil water bicarbonate that did not reach complete isotopic equilibrium with soil atmosphere CO₂ before precipitation. Apparent ages may also be too young owing to contamination by soil carbonate precipitated later. The reliability of ¹⁴C ages of carbonate nodules in similar sections in loess in the semiarid zone of Tunisia has been discussed³ and tentatively confirmed by thermoluminescence dates⁴. The combination of dating problems probably explains much of the spread of ages for each calcic horizon. For example, the younger ages (especially of CH-II) for sections that were cut by wadis may result from contamination by younger carbonate. Alternatively, it may be that the generally fluvial sediments of such sections were in each phase deposited later than the aeolian sediments of other sections. We suggest the following as the best estimates of the true ages of the calcic horizons: ~13,000 yr BP for CH-I, ~28,000 yr BP for CH-II, and ~37,000 yr BP for CH-III. It should be noted that these are estimates of the average ages of the soil nodules; the duration of the period of their formation is not known.

Independent support for the accuracy of these age estimates comes from ^{14}C analyses of land snail shells (reported here corrected for fractionation and age anomaly⁵) and archaeological sites contained within some of the calcic horizons. These date the sediments in which the calcic horizons were formed, not the time of pedogenesis, and therefore must be slightly older than the nodule ages. A snail date from CH-I ($14,600 \pm 500$ yr BP; *Trochoidea seetzeni*) agrees well with the estimated nodule age. CH-I is associated with tools of Epipalaeolithic age (10,000–16,000 yr BP) at one site⁶. At another site (Givat Hayil; local grid coordinates 1178,0425) a weak calcic horizon, presumably CH-I, occurs in sandy sediments which contain an *in situ* Mushabian tool assemblage (excavated and identified by S. Rosen, personal communication). This culture is dated to 13,000–14,000 yr BP⁷. At the confluence of Nahal Beer Sheva and Nahal Besor, the upper calcic horizon (presumably CH-I) overlies an *in situ* Geometric Kebaran site (culture dated to ~14,000 yr BP⁷); a pre-pottery Neolithic B site (culture dated to 9,300–8,000 yr BP⁸) occurs on the surface of the section⁹. All evidence thus points to the reasonable accuracy of the soil nodule dates for CH-I. No cultural artefacts have been found associated with CH-II or CH-III at any of the sites. Because the age of CH-III is near the limit of ^{14}C accuracy, it must be taken as tentative: it could represent an older calcic horizon that is always slightly contaminated.

In sections consisting of fluviially redeposited loess, double calcic horizons sometimes occur. This is seen at Ramat Hovav, where both palaeosol I and II contain two calcic horizons¹⁰. The apparent age of the upper horizon is in both cases younger than the lower, by ~1,000–3,000 yr¹⁰. At Omer, palaeosol II has a double calcic horizon, which merges into a single calcic horizon away from the wadi. In the fluvial sections, sedimentation seems to have occurred discontinuously during the period of calcic horizon formation. In the field, CH-II can usually be distinguished from CH-I and CH-III by its better-developed nodules (5–15 mm versus ~3–5 mm), except at Netivot where it is weakly developed². Palaeosol II has a higher clay content than I and III and consequently develops extensive cracks, giving it a blocky structure.

Modern, well-developed calcic horizons do not occur in the Negev, but are known from the northern coastal plain of Israel¹¹, where mean annual rainfall is ~600 mm. Calcic horizons generally less developed than those described here are also known from colluvium in the northern Negev dating to the early Holocene^{10,12}, which was a significantly wetter period than now¹³. The much reduced pedogenic activity in the Negev today is the result of low rainfall, which produces sparse plant cover

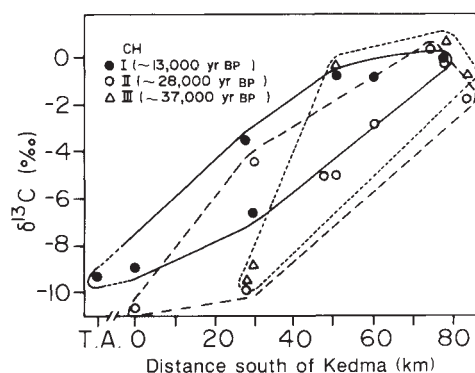


Fig. 2 $\delta^{13}\text{C}$ values of soil nodule carbonate (mean of five nodules, analysed separately) from calcic horizons I, II and III from the study sites. Sites are plotted in relation to their distance south of Kedma; Talme El'azar (T.A.) is plotted separately.

and low plant productivity and therefore little soil CO_2 to dissolve carbonates. The periods of calcic horizon formation in the Negev therefore represent conditions much wetter than now, probably at least as wet as present-day conditions on the coastal plain. Thus three major wet periods, centred on the ^{14}C ages cited above, are indicated for the latter half of the last glacial period in southern Israel. These wet periods must have been separated by drier periods, during which little pedogenic activity occurred. The widespread truncation of palaeosol II^{2,9,10} indicates extensive erosion after formation of CH-II and before accumulation of palaeosol I. This period represents the last glacial maximum centred at ~18,000 yr BP. This extensive erosion probably resulted from the reduction of plant cover due to decreased rainfall. Extensive dune movements⁹ and a reduction in arboreal pollen¹⁴ in other areas of the southern Levant at this time support this interpretation.

Additional evidence of the climatic conditions represented by the calcic horizons comes from analysis of ^{13}C of the nodule carbonate. The ^{13}C content of these carbonates reflects that of the soil atmosphere CO_2 , with which the bicarbonate in soil water undergoes isotopic exchange before carbonate precipitation. As plant productivity increases (with increasing rainfall¹⁵), so soil CO_2 ^{13}C decreases, as plant-derived CO_2 is depleted in ^{13}C . High ^{13}C may also be produced under dry conditions owing to the presence of C_4 plants^{16,17} (here mostly grasses and Chenopodiaceae, the latter being characteristic of drier areas¹⁸). For CH-I, a regular trend of decreasing $\delta^{13}\text{C}$ values is seen from south ($\delta^{13}\text{C} = 0\%$) to north (-10%) (Fig. 2). This implies that a north-south rainfall gradient also existed in the region during the wet period ~13,000 yr ago. CH-II shows a parallel trend in ^{13}C , also indicating a rainfall gradient, but has ^{13}C values more depleted by ~2‰. Thus, wetter conditions are indicated at ~28,000 yr BP than at 13,000 yr BP. CH-III shows depleted values similar to CH-II in the north but more enriched values (as CH-I) in the south; a stronger rainfall gradient that at CH-I or CH-II times is thus implied. The climatic trends indicated by analysis of the palaeosols are summarized in Fig. 3. These results are in agreement with other lines of evidence in the region. In northern Sinai, the CH-II period is represented by extensive sedimentation in wadis¹⁹ and CH-I by the existence of a lake²⁰. Pollen evidence from the central Negev indicates the occurrence of wet periods during the Upper Palaeolithic and late Epipalaeolithic (Natufian)¹⁴, which probably correspond with the periods of formation of CH-II and CH-I, respectively.

Although major climatic fluctuations are indicated for the latter half of the last glacial period in the northern Negev Desert, the existence of a strong rainfall gradient throughout this period implies that the northern boundary of the great desert belt centred at ~23° N (which includes the Saharan and Arabian

Table 1 ^{14}C chronology of calcic horizons I–III

Site	^{14}C age (yr BP)		
	CH-I	CH-II	CH-III
Talme El'azar*	12,650 ± 390		
Kedma*	14,870 ± 460		
Netivot†	13,630 ± 100		35,500 ± 1,500
Magalit‡§		23,300 ± 750	
Omer§		23,500 ± 600	
Beer Sheva	13,500 ± 270	30,700 ± 1,000	34,300 ± 2,600
Ramat Hovav†§	11,680 ± 140	25,400 ± 400	
Besor 2§		26,700 ± 900	
Besor 1	(19,500 ± 450)	30,900 ± 1,400	38,600 ± 4,200
Lavan§		29,200 ± 700	36,800 ± 1,500
estimated age	~13,000 yr BP	~28,000 yr BP	~37,000 yr BP

Sites are arranged from north to south.

* Data from ref. 11.

† Data from ref. 10.

‡ Section described in ref. 6.

§ Section cut by wadi and generally consisting of fluviially redeposited loess.

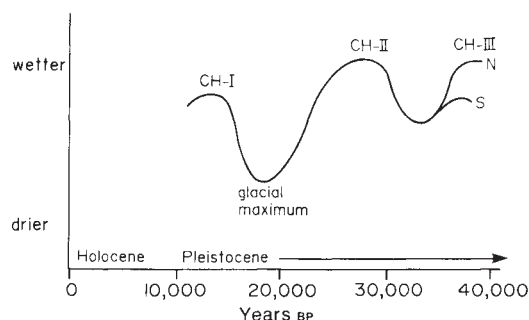


Fig. 3 Qualitative trends in rainfall in the northern Negev Desert during the latter half of the last glacial period, based on analysis of the palaeosols. The duration of the various climatic phases and the rate of change from one to another are not known in detail. N and S designate northern and southern parts of the study area.

deserts) was always in this general region. Last glacial palaeosols are scarce in the central Negev (30° 30' N) and unknown from the southern Negev. Assuming that, as today, well-developed palaeosols form only with ≥ 500 mm mean annual rainfall, this pattern constrains shifts of the present-day isohyets to ≤ 150 km to the south during the wet periods, which is not far on a global scale. Given the lower temperatures of the late Pleistocene and consequent decrease in evapotranspiration, a smaller amount of rainfall would produce equivalent soil moisture conditions, so the shift may have been even less than this. The existence of strong north-south rainfall gradients also implies that, as today, the major source area of precipitation is in the north or north-west; the wet periods were not the result of the northward penetration of monsoonal rains from the Indian Ocean (which today sometimes reach the southern Negev)²¹.

Also along the northern boundary of the Saharan-Arabian desert belt, but far to the west, the loess in southern Tunisia shows a sequence similar to that of the Negev, with a well-developed palaeosol dating to $\sim 28,000$ yr BP (corresponding to CH-II in the Negev) and a more weakly developed one at $\sim 11,000$ yr BP (probably slightly younger than CH-I)³. But a palaeosol dating to $\sim 20,000$ yr BP (a dry period in the Negev) also occurs in this sequence³. This difference implies a more southerly position of the high-pressure cell in north-west Africa than in the Middle East at this time²². Other evidence for the Maghreb region (W. Africa, north of the Sahara), recently reviewed by Rognon²³, also points to a wet period at $\sim 28,000$ yr BP and also to an earlier wet phase, dating from $\sim 40,000$ yr BP (and perhaps synchronous with formation of CH-III in the Negev).

The wetter periods in the Negev appear to correspond with the warmer periods of the last glacial period on a global scale. Interstadials occurred in both North America and Europe at $\sim 38,000$ – $41,000$ yr BP (around the time of formation of CH-III) and at $\sim 28,000$ – $30,000$ yr BP²⁴ (corresponding to CH-II). A high sea stand at $\sim 40,000$ yr BP is well documented and there is limited evidence for a later one at $28,000$ yr BP^{25,26}. CH-I correlates with the Allerød phase—a period of warming following the last glacial maximum, just before the climatic reversal (cooling) of the Younger Dryas²⁷. Thus there appears to be a connection between temperate-zone warming events and southward shifts in the position of the northern boundary of the desert belt in the Middle East.

This work was supported in part by a grant from the Minerva Foundation. We thank the Electric Company of Israel and Dr R. Gerson for providing samples from one of the sites; Mr I. Carmi for providing the radiocarbon dates; Dr A. Rosen for bringing the Magalit section to our attention; T. Galin for assistance with the sample preparation; and Mrs R. Silanikov for assistance with the mass spectrometry.

Received 6 August 1987; accepted 7 January 1988.

1. Yaalon, D. H. & Dan, J. Z. *Geomorph.*, NS 20, 91–105 (1974).
2. Bruins, H. J. & Yaalon, D. H. *Acta geol. acad. Sci. Hung.* 22, 161–169 (1979).
3. Coudé-Gaussen, G., Olive, P. & Rognon, P. *Rev. Géol. dynam. Géogr. phys.* 24, 61–73 (1983).
4. Coudé-Gaussen, G., Rognon, P., Rapp, A. & Nihlén, T. Z. *Geomorph.*, NS 31, 129–144 (1987).
5. Goodfriend, G. A. *Radiocarbon* 29, 159–167 (1987).
6. Rosen, A. *Rep. geol. Surv. Israel* 25/86, 1–38 (1986).
7. Weinstein, J. M. *Radiocarbon* 26, 297–366 (1984).
8. Bar-Yosef, O. *Curr. Anthropol.* 27, 157–162 (1986).
9. Goldberg, P. *Geochronology* 1, 225–244 (1986).
10. Magaritz, M. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 53, 213–229 (1986).
11. Magaritz, M., Kaufman, A. & Yaalon, D. H. *Geoderma* 25, 157–172 (1981).
12. Goodfriend, G. A. *Quat. Res.* 28, 374–392 (1987).
13. Magaritz, M. & Goodfriend, G. A. in *Abrupt Climatic Change—Evidence and Implications* (eds Berger, W. H. & Labeyrie, L. D.) 173–183 (Reidel, Dordrecht, 1987).
14. Horowitz, A. *The Quaternary of Israel* (Academic, New York, 1979).
15. Pittock, A. B. & Nix, H. A. *Climatic Change* 8, 243–255 (1986).
16. Rightmire, C. T. & Hanshaw, B. B. *Wat. Resources Res.* 9, 958–967 (1973).
17. Cerling, T. E. *Earth planet. Sci. Lett.* 71, 229–240 (1984).
18. Shomer-Ilan, A., Nissenbaum, A. & Waisel, Y. *Oecologia* 48, 244–248 (1981).
19. Goldberg, P. Z. *Geomorph.*, NS 28, 193–217 (1984).
20. Goldberg, P. in *Prehistoric Investigations in Gebel Maghara, Northern Sinai* (eds Bar-Yosef, O. & Phillips, J. L.) *Qedem* (7), 11–31 (1977).
21. Gat, J. R. & Carmi, I. in *The Influence of Climate Change and Climatic Variability on Hydrological Regime and Water Resources* (eds Solomon, S. I., Beran, M. & Hogg, W.) 513–523 (IAHS, Wallingford, 1987).
22. Rognon, P. *Rev. Géol. dynam. Géogr. phys.* 22, 313–328 (1980).
23. Rognon, P. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 58, 11–34 (1987).
24. Stuiver, M., Heusser, C. J. & Yang, I. C. *Science* 200, 16–21 (1978).
25. Moore, W. S. in *Uranium Series Disequilibrium: Applications to Environmental Problems* (eds Ivanovich, M. & Harmon, R. S.) 481–496 (Clarendon, Oxford, 1982).
26. Chappell, J. & Shackleton, N. J. *Nature* 324, 137–140 (1986).
27. Berger, W. H., Killingley, J. S., Metzler, C. V. & Vincent, E. *Quat. Res.* 23, 258–271 (1985).

Rhodium distribution at the Cretaceous/Tertiary boundary analysed by ultrasensitive laser photoionization

G. I. Bekov*, V. S. Letokhov*, V. N. Radaev*,
D. D. Badyukov† & M. A. Nazarov†

* Institute of Spectroscopy, USSR Academy of Sciences,
142092 Troitsk, Moscow Region, USSR

† Vernadsky Institute of Geochemistry and Analytical Chemistry,
USSR Academy of Sciences, 117975 Moscow, USSR

The extraordinarily high concentrations of iridium and other siderophiles found in Cretaceous/Tertiary (K/T) boundary deposits have been interpreted¹ as a result of a large extraterrestrial body falling upon the Earth, causing mass extinction of the biota dominant in the Cretaceous. The existing alternative models of the Cretaceous terminal event try to explain the anomaly of siderophile elements by their being concentrated through sedimentation or volcanic eruptions^{2,3}. To understand the nature of the K/T event, it is important to establish the proportions of siderophiles and their isotopes in boundary deposits of that age, especially the proportions of platinum-group elements, their concentrations in extraterrestrial material being high and occurring in ratios differing from those typical of terrestrial material. Here we present the first data on rhodium concentrations at the K/T boundary in the Sumbar-SM-4 section (Turkmen SSR) obtained by the ultrasensitive laser photoionization spectroscopy (LAPIS) technique^{4,5}. The maximum Rh concentration in the samples studied is 24.2 ng g^{-1} . The Rh/Ir ratio is 0.34 ± 0.06 , which is close to the cosmic ratio of these elements. The LAPIS technique is efficient and can be used for determining the concentrations of all siderophile elements and their isotopes in geological materials.

According to the impact theory, the extraterrestrial body that fell upon the Earth more than 65 Myr ago could have had a chondritic^{1,6}, iron^{7,8} or cometary⁹ composition, its material entering the K/T boundary sediments together with the material of crater ejecta, as borne out by the findings of shocked quartz grains^{10,11} in these sediments. Extensive studies of the K/T boundary geological materials have so far revealed an iridium anomaly¹² in a few tens of such geological sections, but only in some of them were the concentrations of Pt, Pd, Os and Ru

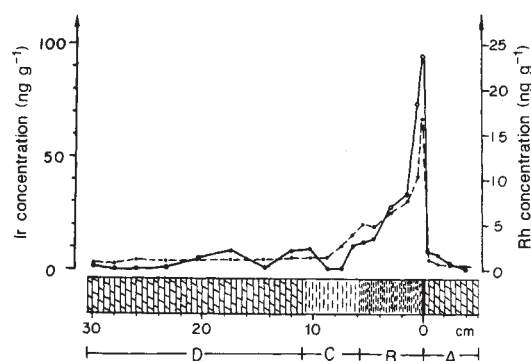


Fig. 1 Distributions of Ir (broken line, left-hand scale) and Rh (solid line, right-hand scale) in the Sumbar-SM-4 section (Turkmen SSR). In this section the Maastrichtian deposits are represented by marl (layer A) containing abundant ammonites up to the Danian boundary. The Danian sediments start with a brown clay layer 4–6 cm in thickness (layer B), the basal portions of which (1–2 cm) are highly enriched with gypsum and limonite. Observed here and there at the base of this layer are lenses and pockets of a yellowish clay material also exceptionally rich in gypsum and ferrous and ferric hydroxides. Higher up the section, the brown clay is gradually replaced by a greenish-grey clay (layer C) 6–10 cm thick. Still higher, this layer gradually gives way to marl alternating with limestone and clay to form the deposits of the Sumbar horizon (layer D).

measured^{6,7,13}. This difference in the number of measurements between iridium and the other platinum-group elements is explained by the fact that the neutron-activation analysis technique used in most measurements shows the highest sensitivity to Ir only.

At the same time, the most reliable inferences about the nature of such anomalies can be drawn not so much from the absolute concentrations of elements, as from their relative concentrations in geological samples. An important characteristic of such objects is the proportion in them of the isotopes of siderophile elements, as demonstrated, for example, for Os (ref. 14). The LAPIS technique is promising for such investigations, as it allows the measurement of the concentrations of platinum-group elements and, in principle, of their isotopes in geological objects with a sensitivity of up to 1 p.p.t., which is substantially below the background level of these elements under terrestrial conditions. We used this technique to determine the Rh content of the samples from the Sumbar-SM-4 section (Fig. 1), in which the concentrations of Ir and other elements had already been determined^{8,15}. The samples 1–2 cm thick were taken continuously and represented a ~40-cm interval of K/T boundary deposits.

Geological samples were analysed both directly and after chemical pretreatment. Acid-alkali dissolution of samples, followed by the sorption of rhodium on a polymeric thioether¹⁶ or ferric hydroxide, proved most suitable. The sample under analysis was placed in a pyrolytic graphite crucible and atomized *in vacuo* by being heated in a stepwise fashion to 2,300 °C. The atomic-molecular beam thus formed was irradiated with three dye lasers to effect selective three-step photoionization of rhodium atoms (Fig. 2). To improve the selectivity of the analysis, the ions were mass-separated. The laser photoionization analysis technique is described in detail in refs 4, 5 and 17. The detection limit for Rh in the pretreated samples was 3 pg g⁻¹. The blank values did not exceed 30 pg g⁻¹.

Figure 1 shows the results of the Rh determination together with data on Ir borrowed from ref. 15. The maximum Rh content (24.2 ng g⁻¹) was found in a pocket of a yellow clay material directly at the K/T boundary. Higher up the section the concentration of Rh gradually fell to 0.5 ng g⁻¹ at 23 cm from the boundary and remained at this level at least to 40 cm from the boundary. Downwards from the boundary the Rh content

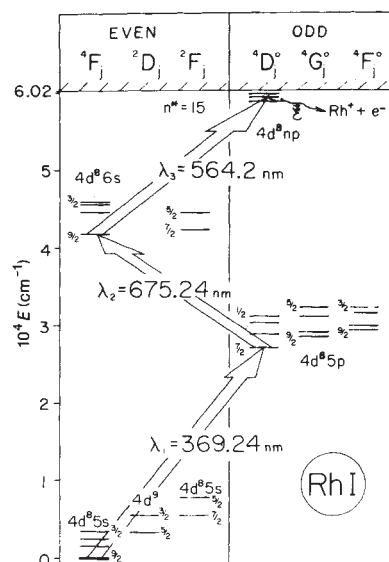


Fig. 2 Energy-level diagram of Rh atom and laser transitions used for its ionization. Three concurrent dye-laser pulses raise ~10% of Rh atoms to a Rydberg state with an effective principal quantum number of $n^* \approx 15$; 20 ns later the highly excited atoms are ionized with an efficiency close to unity by means of an electric field pulse.

decreased rapidly to 0.3 ng g⁻¹ at 4 cm below the boundary. At 80 cm below the boundary the Rh concentration was 0.07 ng g⁻¹, which can probably be taken to be the background abundance. The concentrations of Rh and Ir correlate well ($r = 0.96$). The Rh/Ir ratio is in the range 0.1–1.5, its average value being 0.47 ± 0.08 (1 σ). The integrated amount of Rh in the Sumbar-SM-4 section is estimated at 200 ng cm⁻². With the total amount of Ir in this section being 580 ng cm⁻² (ref. 15), the integrated Rh/Ir ratio will be 0.34, which is below average and hence points to the possibility of some low-rhodium samples being contaminated with Ir. Nevertheless, the closeness of the integral and average Rh/Ir ratios, along with the strong correlation between these elements, gives reason to believe that the Rh/Ir ratio observed was determined by the composition of the source of the K/T siderophiles, and not by fractionation processes that occurred in the course of sedimentation.

Our results indicate that Rh, as well as the other platinum-group elements, clearly marks the K/T boundary. Its concentration in the K/T boundary deposits may be at least two orders of magnitude greater than background (~0.1 ng g⁻¹). Such high or even higher concentrations of Rh in the Earth's crust are known only from ultramafic rocks and Cu–Ni ore deposits, which limits the possibilities of the earthly interpretations of the origin of the K/T boundary anomaly of platinum-group elements. At the same time, the Rh concentrations observed in the K/T boundary deposits can readily be explained by the presence of extraterrestrial material, which on average (chondritic material) contains 150–500 ng g⁻¹ of Rh (refs 18–21), that is, by at least three orders of magnitude more than the rocks of the Earth's crust. The low (0.34) Rh/Ir ratio in the K/T boundary deposits is also consistent with the hypothesis of extraterrestrial origin of the platinum-group elements in them, being practically identical with the chondritic ratio (0.3–0.6 (refs 18–21)). In the Earth's rocks this ratio is usually higher. It comes to 0.7 in mantle modules²², 1.0 in kimberlites²², 11 in ultramafics²³, and 1.4–4 in Cu–Ni ore deposits^{24,25}. Thus, in accordance with the higher mobility of Rh²⁶, the above series points to an increase in Rh/Ir with increasing degree of fractionation of matter, and it can be expected that in volcanic smokes, which were considered the source of platinum-group elements in the K/T boundary deposits², the Rh/Ir ratio must substantially exceed the chondritic value.

On the whole, the Rh/Ir proportions observed at the K/T boundary suggest that the anomalous quantities of platinum-group elements in the K/T boundary sediments did not result from their concentration in the sedimentation cycle and have no relation to volcanic sources. The data obtained point unambiguously to an extraterrestrial origin for the K/T boundary anomaly and bear out the impact theory¹ of the Cretaceous terminal event.

Received 28 July; accepted 8 December 1987.

1. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095–1108 (1980).
2. Zoller, W. H., Parrington, J. R. & Kotra, J. M. P. *Science* **222**, 1118–1120 (1983).
3. Officer, C. B. & Drake, C. L. *Science* **219**, 1383–1390 (1982).
4. Bekov, G. I. & Letokhov, V. S. in *Laser Analytical Spectrochemistry* (ed. Letokhov, V. S.) 98–151 (Adam Hilger, Bristol, 1986).
5. Letokhov, V. S. *Laser Photoionization Spectroscopy* (Academic, New York, 1987).
6. Ganapathy, R. *Science* **209**, 921–923 (1980).
7. Kyte, F. T., Zhou, Z. & Wasson, J. T. *Nature* **283**, 651–656 (1980).
8. Nazarov, M. A., Alekseev, A. S. & Ammaniyazov, K. N. *Lunar planet. Sci.* **17**, 605–606 (1986).
9. Hsü, K. S. *Nature* **285**, 201–203 (1980).
10. Bohor, B. F., Foord, E. E., Modreski, P. I. & Triplehorn, D. M. *Science* **224**, 867–869 (1984).
11. Badjukov, D. D., Nazarov, M. A. & Suponeva, I. V. *Lunar planet. Sci.* **17**, 18–19 (1986).
12. Alvarez, W., Alvarez, L. W., Asaro, F. & Michel, H. V. *Spec. Pap. geol. Soc. Am.* **190**, 305–315 (1982).
13. Kyte, F. T., Smit, J. & Wasson, J. T. *Earth planet. Sci. Lett.* **73**, 183–195 (1985).
14. Lichte, F. E. *et al. Nature* **322**, 816–817 (1986).
15. Alekseev, A. S., Barsukova, L. D., Kolesov, G. M., Nazarov, M. A. & Ammaniyazov, K. N. *Lunar planet. Sci.* **17**, 9–10 (1986).
16. Petrukhin, O. M., Malofeeva, G. I., Nefedov, V. I. & Salyn, J. V. *Zh. analyt. Khim.* **38**, 250–255 (1983).
17. Bekov, G. I., Letokhov, V. S. & Radaev, V. N. *J. opt. Soc. Am. B* **2**, 1554–1560 (1985).
18. Sen Gupta, J. G. *Analytica chim. Acta* **42**, 481–488 (1968).
19. Sen Gupta, J. G. *Chem. Geol.* **3**, 293–305 (1968).
20. Anders, E. & Ebihara, M. *Geochim. cosmochim. Acta* **46**, 2363–2380 (1982).
21. Wasson, J. T. *Meteorites: Their Record of Early Solar-System History* (Freeman, New York, 1985).
22. Kaminskii, F. V., Frantesson, E. V. & Khvostova, V. P. *Dokl. Akad. Nauk SSSR* **219**, 204–207 (1974).
23. Hiemstra, S. A. *Econ. Geol.* **80**, 944–957 (1985).
24. Cowden, A., Donaldson, M. J., Naldrett, A. J. & Campbell, I. H. *Can. Miner.* **23**, 300 (1985).
25. Naldrett, A. J. & Duke, J. M. *Science* **208**, 1417–1424 (1980).
26. Cousins, C. A. *Trans. geol. Soc. S. Afr.* **76**, 77–81 (1973).

Biotransformations of organosulphur compounds in sediments via 3-mercaptopropionate

Ronald P. Kiene & Barrie F. Taylor

Department of Marine and Atmospheric Chemistry,
Rosenstiel School of Marine and Atmospheric Sciences,
University of Miami, 4600 Rickenbacker Causeway, Miami
Florida 33149-1098, USA

3-Mercaptopropionate (3-MPA) is a major organic sulphur compound (thiol) recently detected in the anoxic pore waters of coastal marine sediments¹. A chemical mechanism for the formation of 3-MPA has been identified as the abiotic addition of H₂S to the double bond of acrylate². A natural source of acrylate (for the addition reaction) is dimethylsulphoniopropionate (DMSP) which yields dimethylsulphide (DMS) and acrylate upon enzymatic cleavage^{3,4}. DMSP occurs in plants, particularly those in saline environments where it fulfils an osmotic function⁵. Here we present evidence that 3-MPA is generated from the biotransformations of DMSP as well as the sulphur-containing amino acids methionine and homocysteine. Additionally, 3-MPA is metabolized in sediments. In view of these findings we suggest that 3-MPA is a central metabolite in both catabolic and assimilatory sulphur metabolism in general.

Sediment-core samples were obtained from Bear Cut in Biscayne Bay, Florida, using polybutyrate core tubes. Depth profiles were obtained by side-port sampling with cutoff polyethylene syringes. All core manipulations were carried out in an anoxic chamber. Anoxic sediment slurries were prepared from the upper 15 cm of Bear Cut sediments using previously published procedures⁴. Sediment samples were centrifuged to obtain pore

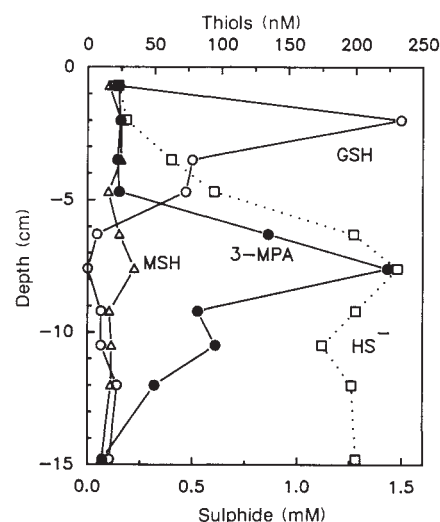


Fig. 1 Depth profiles of dissolved thiols and sulphide (sulphide denoted by $\square$) in the pore waters of a sediment from Bear Cut, Key Biscayne, Florida. 3-MPA ($\bullet$) is 3-mercaptopropionate, GSH ($\circ$) is glutathione and MSH (Δ) is methanethiol. Sediment samples were taken with cut-off syringes from pre-drilled holes in the sides of the core tube. Whole sediment was centrifuged in Centrux filter units using a 0.45 μ m pore-size Teflon filter. Pore water was then analysed for thiols and sulphide. All manipulations were carried out in an anoxic chamber. The detection limits for thiols were 3, 4 and 10 nM for 3-MPA, MSH and GSH respectively.

water and immediately derivatized for the analysis of thiols by HPLC⁶. Additions of organic sulphur compounds and inhibitors were made to slurries by injecting aqueous solutions through the rubber stoppers. Sulphide was measured by the method of Cline⁷.

Figure 1 shows a depth profile for sulphide and several dissolved thiols in the pore waters of a sediment collected from Bear Cut near Miami. 3-MPA and glutathione were the principal thiols in the pore waters and they reached similar maximal levels of ~200 nM. However, glutathione occurred nearer to the sediment-water interface, whereas the peak in 3-MPA was deeper and coincided with the highest sulphide concentrations. Methanethiol was present at 10–35 nM throughout the sediment column with a maximum at ~7 cm.

The presence of thiols in sediment pore waters is of interest because of their chemical and biochemical reactivity. Glutathione predominates near the sediment-water interface where it might be associated with aerobic organisms⁸. Previous studies⁹ indicated that glutathione was an important thiol in anoxic sediments. Our data support this suggestion, although the highest levels we have seen in pore waters were only 600 nM, which are much lower than the 2-mM concentrations reported by Luther *et al.*⁹ in *Spartina* salt marsh sediments.

The mechanisms responsible for the formation of 3-MPA in these sediments are not known, although the presence of high levels of sulphide indicate that the acrylate-sulphide addition pathway is possible². However, experiments with sediment slurries revealed that homocysteine and methionine, in addition to DMSP, were also precursors of 3-MPA and methanethiol (Fig. 2). In these experiments the rate of 3-MPA accumulation from acrylate was slow in comparison to other precursors, even though sulphide concentrations exceeded 0.5 mM. This suggests that other mechanisms were involved in the formation of 3-MPA from the organosulphur compounds. 3-MPA was also consumed when it was added to sediment slurries and a biological contribution to the consumption was indicated by the inhibitory effects of antibiotics (Fig. 3), and by autoclaving. The abiological losses are due to the binding or adsorption of thiols to sediment particles¹.

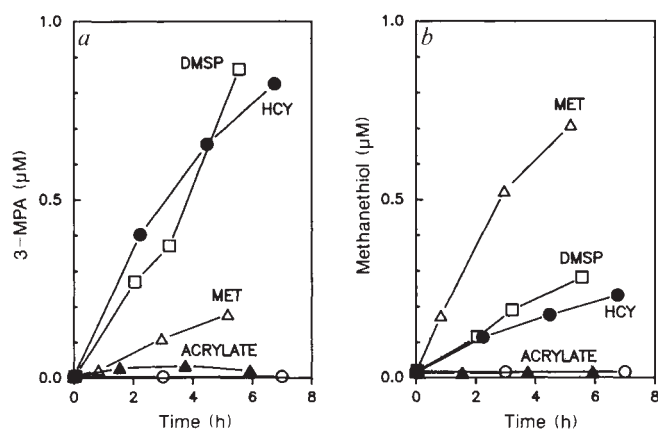


Fig. 2 Production of 3-MPA (a) and methanethiol (b) from DMSP (□), homocysteine (●), methionine (Δ), and acrylic acid (▲) in anoxic sediment slurries prepared from Bear Cut sediments. Samples were pre-incubated for 24 h to allow endogenous levels of 3-MPA to be consumed; then each precursor compound was added at 10 μM. Unamended sediments (○) did not accumulate any additional 3-MPA or methanethiol. Data presented are from single sample bottles. The experiment was repeated several times with similar results.

Our findings draw attention to 3-MPA as an important organic sulphur compound in the sediments, and as a central intermediate in the metabolism of DMSP, methionine and homocysteine. 3-MPA is an inhibitor of fatty acid oxidation when added in mammalian systems¹⁰. Its inhibitory action is due to the formation of 3-mercaptopyruvyl-CoA which, because of its terminal -SH group, substitutes for CoA in some enzymatic reactions¹¹. Our data suggest that, at the concentrations found in the sediments, 3-MPA is metabolized and pore-water pools may therefore turn over. However, the biochemical fate of 3-MPA in these systems remains unknown. The sources of 3-MPA include acrylate and several ubiquitous organic sulphur compounds (DMSP, methionine and homocysteine), and the sinks probably include microbial metabolism, diffusion and chemical adsorption to sediment¹. Production of 3-MPA in oxygenated zones or upward diffusion from anaerobic zones could result in oxidation to the disulphide and further aerobic metabolism. The rate at which 3-MPA turns over in sediments is in need of investigation.

The inter-relationships between the various organic sulphur compounds examined here are summarized in Fig. 4. DMSP may give rise to 3-MPA either after cleavage to DMS and acrylate followed by the chemical addition of H₂S to acrylate, or by successive demethylations with 3-methylpropionate as an intermediate. In this study the demethylation pathway dominated the chemical addition reaction. The rates of 3-MPA formation from acrylate that we observed were similar to those reported by Vairavamurthy and Mopper² for an abiotic system. However, the chemical formation of 3-MPA was slow in comparison with the biological production from similar levels of DMSP. The production of 3-MPA by the chemical addition reaction does, however, illustrate a potentially significant mechanism whereby inorganic sulphur could be incorporated into the organic fraction of sediments². The addition of H₂S to α,β-unsaturated compounds as described by Vairavamurthy and Mopper², may be of greater geochemical importance with more complex molecules where the precursors and the products might not be rapidly metabolized.

Methionine is converted to 3-methylpropionate in animal tissues, with 2-keto-4-methylbutyrate as an intermediate¹². Demethylation of 3-methylpropionate would give 3-MPA. Similarly, homocysteine could be converted to 3-MPA via 2-keto-4-mercaptobutyrate. The identification of homocysteine and methionine as precursors of 3-MPA raises the distinct possibility of the widespread occurrence and importance of

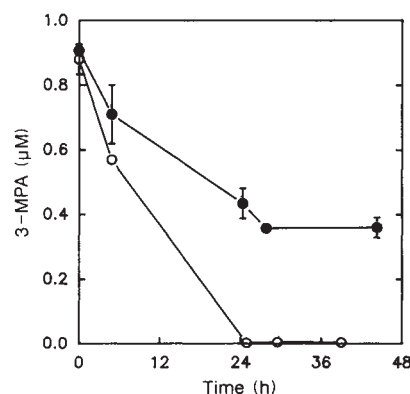


Fig. 3 Time course of 3-MPA in untreated (○) anoxic sediments and those treated with a combination of the antibiotics chloramphenicol (125 μg ml⁻¹) and tetracycline (60 μg ml⁻¹, ●). Sediments were preincubated for 20 h with antibiotics before addition of 1 μM 3-mercaptopyruvate. Data represent the mean of two replicate samples with the error bars indicating the range in the data.

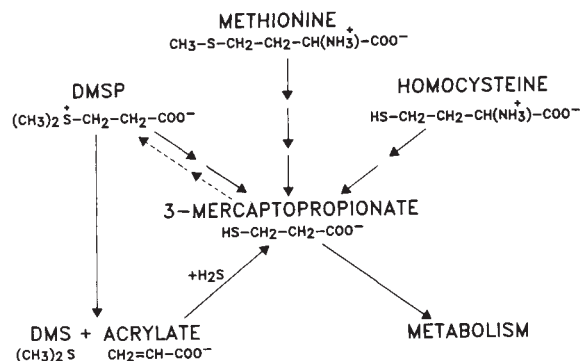


Fig. 4 A proposed scheme of how several important organic sulphur compounds are transformed in natural environments. The central role of 3-mercaptopyruvate in these transformations is illustrated. Solid lines represent conversions established in the present work and dotted lines are speculated pathways. Multiple arrows indicate multistep conversions.

3-MPA in organic sulphur transformations in non-saline as well as saline environments.

The possible use of 3-MPA as an assimilatory source of reduced sulphur is intriguing; methylations could retrieve DMSP for osmotic functions. Evidence for the close coupling of biochemical reactions involving organic sulphur compounds includes observations that methionine can donate both methyl groups, the sulphur atom and the carbon skeleton in the biosynthesis of DMSP in *Ulva lactuca*¹³. Additionally, DMSP may act as a methyl donor to homocysteine in the formation of methionine¹⁴. Homocysteine, in addition to being a precursor for 3-MPA, was converted to methanethiol (Fig. 2b), possibly after methylation to form methionine and subsequent cleavage to methanethiol⁴. Methylation reactions involving reduced sulphur compounds occur in anoxic sediments¹⁵ and have been observed in a variety of bacteria and eucaryotes^{16,17}. A potential methyl donor involved in these reactions is S-adenosylmethionine¹⁶⁻¹⁸, which further suggests the close biochemical coupling of these compounds in natural habitats. Although previously unknown, 3-MPA could play a key role as a common intermediate in the metabolism of organic sulphur compounds. Clearly, further work is needed to define the routes and mechanisms for 3-MPA formation and metabolism in natural environments.

We thank K. Mopper and A. Vairavamurthy for valuable discussions and Kirk Malloy for helping with HPLC analyses. This work was supported by the NSF.

Received 16 September 1987; accepted 14 January 1988.

- Mopper, K. & Taylor, B. F. in *Organic Marine Geochemistry* ACS Symp. Ser. No. 305 (ed. Sohn, M.) 324-339 (Am. Chem. Soc., Washington, DC, 1986).
- Vairavamurthy, A. & Mopper, K. *Nature* **329**, 623-625 (1987).
- Cantoni, G. L. & Anderson, D. G. *J. biol. Chem.* **222**, 171-177 (1956).
- Kiene, R. P. & Visscher, P. T. *Appl. envir. Microbiol.* **53**, 2426-2434 (1987).
- Vairavamurthy, A., Andree, M. O. & Iversen, R. L. *Limnol. Oceanogr.* **30**, 59-70 (1985).
- Mopper, K. & Delmas, D. *Analyt. Chem.* **56**, 2557-2560 (1984).
- Cline, J. D. *Limnol. Oceanogr.* **14**, 454-458 (1969).
- Fahey, R. C., Brown, W. C., Adams, W. B. & Worsham, M. B. *J. Bacteriol.* **133**, 1126-1129 (1978).
- Luther, G. W., Church, T. M., Scudlark, J. R. & Cosman, M. *Science* **232**, 746-749 (1986).

- Sabbagh, E., Cuebas, D. & Schulz, H. *J. biol. Chem.* **260**, 7337-7342 (1985).
- Cuebas, D., Beckmann, J. D., Frerman, F. E. & Schulz, H. *J. biol. Chem.* **260**, 7330-7336 (1985).
- Cooper, A. J. L. *Rev. Biochem.* **52**, 187-222 (1983).
- Greene, R. C. *J. biol. Chem.* **237**, 2251-2254 (1962).
- Durell, J., Anderson, D. G. & Cantoni, G. L. *Biochim. Biophys. Acta* **26**, 270-282 (1957).
- Kiene, R. P. & Capone, D. G. *Microbial Ecol.* (in the press).
- Drotar, A. M., Fall, L. R., Mishalanie, E. A., Tavernier, J. E. & Fall, R. *Appl. envir. Microbiol.* **53**, 2111-2118 (1987).
- Drotar, A. M., Burton, G. A., Tavernier, J. E. & Fall, R. *Appl. envir. Microbiol.* **53**, 1626-1631 (1987).
- Maw, G. A. & du Vigneaud, V. *J. biol. Chem.* **176**, 1037-1045 (1948).

Early cultural evidence from Monte Verde in Chile

Tom D. Dillehay* & Michael B. Collins†

* Department of Anthropology, University of Kentucky, Lexington, Kentucky 40504, USA and Universidad Austral de Chile, Casilla 567, Valdivia, Chile

† 7701 North Lamar, Suite 104, Austin, Texas 78752, USA

The entry of the first Asians into the New World is generally thought to have occurred no earlier than 12,000 years ago^{1,2}. Recent archaeological evidence from South America suggests that the migration from Asia to North America might have taken place much earlier. This evidence comes from the Brazilian site of Boqueirão do Sitio da Pedra Furada^{3,4}, with a long cultural sequence possibly extending as far back as 32,000 yr BP, and the Chilean site of Monte Verde^{5,6}. This latter site has one well-documented cultural episode radiocarbon dated at 13,000 yr BP⁷ and another possible one at 33,000 yr BP. We report here two carbon-14 dates from charcoal taken from cultural features associated with the older materials of ~33,000 yr BP. These findings provide additional evidence that people colonized the Americas much earlier than was previously thought.

The Chilean site at Monte Verde is shown in Fig. 1. Since 1977, we have concentrated our research at Monte Verde on the 13,000 year old culture which is characterized by bone and stone technologies (Fig. 2a-c) and by well-preserved wood artefacts, dwelling foundations of both earth and wood, and abundant remains of useful plants, all revealing an early lifeway more culturally diversified than archaeologists had previously demon-

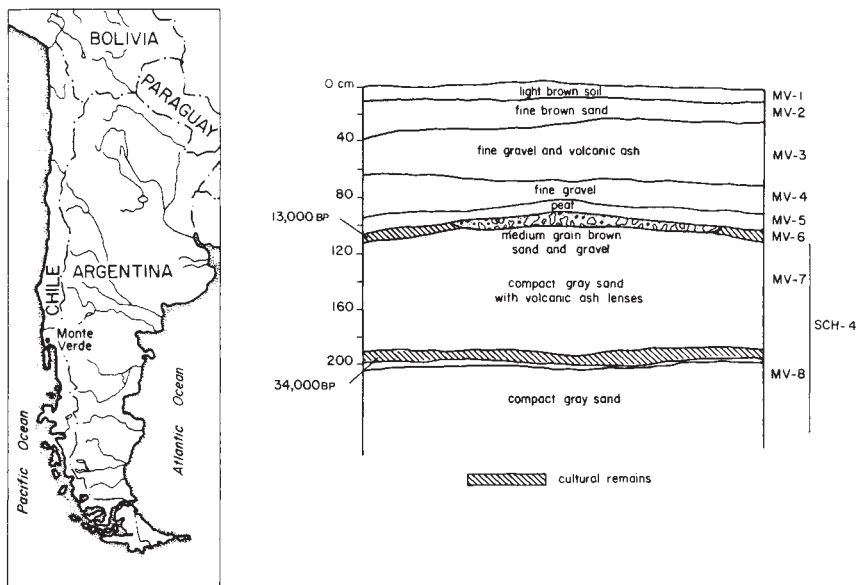
strated. In recent seasons, we have excavated a deeper stratigraphic layer containing 26 lithics in direct association with three apparent cultural features.

Cultural materials at Monte Verde are buried in the banks and adjacent sandy terraces of Chinchihuapi Creek, a southern tributary of the Rio Maullin, located in south-central Chile. The stream drains a wet boggy area in a temperate rainforest that has existed there since at least late Pleistocene times. The site was discovered by a team of investigators from the Universidad Austral de Chile, Valdivia.

The 26 lithics recovered from the buried embankment of an old stream have been classified into five categories: edge-battered ($N=2$), cores ($N=1$), flakes ($N=4$), percussion-split pebble ($N=1$), and naturally fractured but culturally used stones ($N=18$). Eleven stones show clear cultural modification by percussion (Fig. 2d, e) and/or by use of their sharp edges (Fig. 2f, g). Fifteen specimens, which are naturally fractured and exhibit sharp edges and acute angles (15-40°) suitable for cutting or scraping, reveal no clear evidence of alteration by human agency. But their direct association with the hearth-like features and other demonstrable cultural stones leads us to suspect that they were collected and transported to the site by humans. Most of the stones are of basalt and andesite, though a few of tonalite and diorite occur as well.

The two edge-battered stones, one made of basalt and the other of tonalite (Fig. 2d), show clear battering on one edge. The one core is a percussion-struck basalt cobble with a prepared platform and at least ten flake scars on the core face (Fig. 2e). One percussion-split basalt pebble was also found (Fig. 4g). Four basalt flakes show deliberate flaking on their dorsal faces (Fig. 2f). Two of these exhibit macroscopic edge damage in the form of scalar and stepped fractures and a series of scars and striations lying diagonal and perpendicular to the edge. The

Fig. 1 The site of Monte Verde is located in south-central Chile. The geology of the site consists of two stratigraphic units⁸, the Salto Chico unit and the Monte Verde unit, comprised of eight strata. (The drawing is a composite stratigraphic profile of the two horizontally separate and indirectly superimposed cultural components.) The lowest is the Salto Chico Unit (SCH-4) made up of a bluish grey sand and thin lenses of small gravel (MV-8) overlaid by grey coarse-grained sand with fine lenses of volcanic ash (MV-7). The older materials and features lie at the base of the grey sand with thin ash lenses (MV-7) of SCH-4. The Monte Verde unit (MV-1 to MV-6) superimposes the SCH-4 unit. Overlying a narrow strip of SCH-4 (MV-7), is MV-6, a 4-6 m wide ancient creek bed deposit, consisting of medium-grained brown sand and gravels. The 13,000 yr BP cultural materials at Monte Verde rest on the shoreline of the old creek channel (MV-6) and on the sandy banks of that creek (upper MV-7). The 33,000 yr BP remains were recovered from the bank of an older, smaller stream located about 1.8 m below and 50 m south of the present day channel of Chinchihuapi Creek.



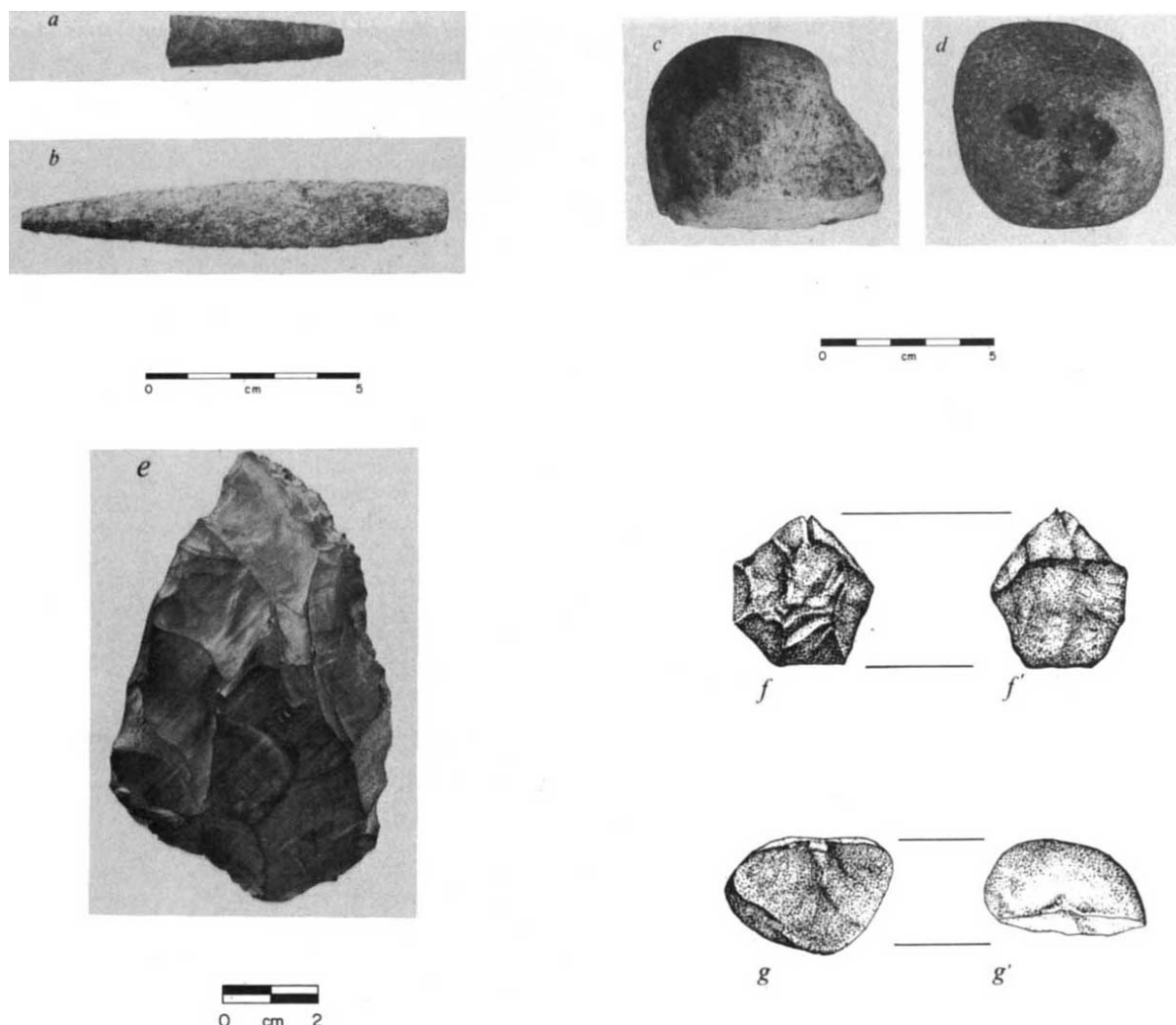


Fig. 2 13,000 yr BP level: *a*, broken and *b*, unbroken bifacially chipped tools made of granitoid; *c*, a large, burned core made of diorite. 33,000 yr BP level: *d*, an edge-battered cobble made of basalt; *e*, a unifacially flaked core made of a fine grained basalt; *f*, percussion-split pebble made of basalt; *g*, secondary percussion flake made of basalt.

remaining 18 pieces are naturally split stones, probably scavenged from distant gravel deposits and brought to the site for use. Two of these specimens show clear evidence of micro-use wear as evidenced by long, shallow, parallel striations and polish indicative of scraping or cutting plant material or fresh wood. An extensive series of facets is visible on the edge of another stone.

Although this older collection of stones has no cultural affinities with assemblages from other known sites, it has some similarities with the 13,000 yr BP lithics. Both collections are dominated by naturally fractured stones with sharp, acute edges which were culturally selected and used. (Our evidence suggests that the Monte Verdeans usually did little more than pick up conveniently fractured basalts and andesites in the nearby creek bed and use them without deliberate chipping or retouch.) The edge angles and the location of worked aspects along the edge of used specimens are similar in both collections.

Two concentrations of charcoal specks in hearth-like basins (H-S 1 and 2) and one clay-lined feature (H-S 3) were documented in direct association with the older lithics. These features were located in small, shallow, roughly circular basins ranging in depth from 2 to 3 cm and in diameter from 40 to 70–80 cm (Fig. 3). Features H-S 1 and 2 were defined by ash and specks of charcoal scattered in and around them and by burned sand

underneath them. Feature H-S 3 is a roughly oval, partially disturbed, clay-lined basin containing specks of burned charcoal (*Nothofagus* sp.) and unburned and partially burned *Juncus* sp., an edible reed. Flotation of all feature fill yielded a small quantity of charred and uncharred seeds and root fragments of *Juncus* sp. The absence of burned rock in these features is not unexpected. Few stones were recovered from the numerous shallow, ash and wood- or clay-lined hearths and braziers in the 13,000 yr BP cultural level. The recovery of both reed and beech vegetation suggests that the climate of 33,000 yr BP was probably cooler and wetter than that of the present day and that the environment was characterized by open forests and marshes, lagoons, and rivers.

There is no evidence to suggest that the features and stones were redeposited. Also, the cultural materials apparently did not filter down through verticracks from the 13,000 yr BP cultural level because the closest materials from the younger occupation are buried about 50 m to the north on the other side of the ancient creek channel (Fig. 4). Most of the stones were concentrated around the features in a 200 m² area of the ancient stream bank. Similar features were not observed on the same stratigraphic surface in 55 other test pits and excavation blocks (constituting ~250 m² of exposed surface) placed in the site area and in non-cultural areas well beyond the vicinity of the

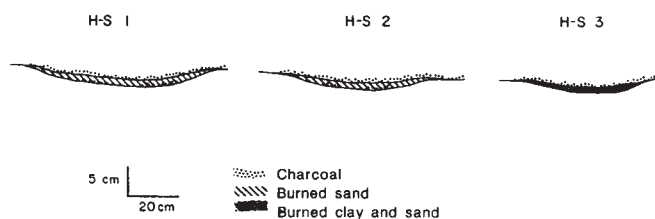


Fig. 3 Profile of the three features in the site.

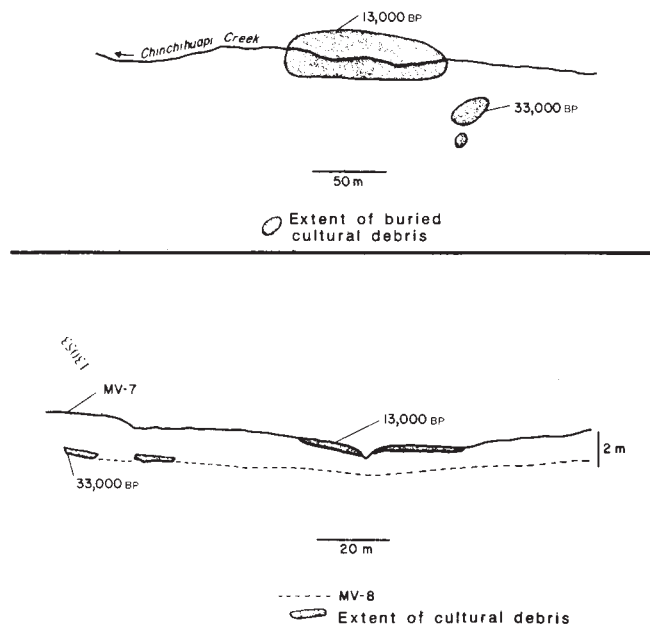


Fig. 4 Vertical and horizontal location of the 13,000 and 33,000 yr BP components in the site.

site. Nor was there any evidence to suggest that these features are of shallow depressions where concentrations of charcoal and organic debris were deposited below an erosional unconformity. One sample of charcoal from H-S 1 and one sample of charred wood from H-S 2 were dated at 33,370 $\pm$ 530 yr BP (β -6724) and $\gg$ 33,020 BP (β -7825), respectively. The latter date was derived from a small portion (0.6 g) of burned wood which resulted in an unresolvable age which could range between 33,020 and 40,000 yr BP (J. Stipp, personal communication). There is no sign of contamination of the radiocarbon materials, and they are in stratigraphic agreement with 16 other radiometric ages associated with the entire sequence in the site. (The volcanic ash lenses in the MV-7 and MV-8 strata were not dated because they were redeposited in the site from older contexts.)

There is no disturbed or reversed stratigraphy or any phenomenon that might account for either the horizontal concentration and exclusivity of the lithics and features on the bank of an ancient stream or other body of water, or their vertical placement within a 5–10 cm thick layer and their horizontal distribution within a 200 m² area (Fig. 4) at the base of the MV-7 in the deeper SCH-4 unit. The tight stratigraphic association and patterning of the stone assemblage around the three features and the absence of stones in other exposed areas of the base of MV-7 suggest that they were not redeposited in this unit.

At present, there is no evidence to explain the 20,000 yr cultural hiatus between these two episodes. Geological studies in the region show that the 56,000–13,000 yr BP period was characterized by glacial activity, changing climatic conditions, and particularly long cold periods^{9,10}. An interstadial character-

ized by a warmer but still cool, wet climate might have occurred between 37,000 and 33,000 yr BP¹¹. After this period, cold conditions continued until deglaciation began by 15,000–14,500 yr BP. It is possible that the climate impeded human occupation of the area during this hiatus or simply that only two human groups ever chose to settle the site area.

Caution should still be taken in considering an earlier date of entry of Asians into the New World. Although the earlier cultural evidence from the site of Monte Verde should be taken as a solid contender for an entry possibly as far back as 30,000 yr BP, we need more information from sites like Monte Verde and Pedra Furada for further evaluation of a human presence in South America much longer than 13,000 years ago.

Received 27 October 1987; accepted 14 January 1988.

1. Bray, W. *Nature* **321**, 726 (1986).
2. Dincauze, D. *Advances in World Archaeology* (eds. Wendorf, F. & Close, A.) 275–323 (Academic, London, 1984).
3. Guidon, N. & Delibrias, G. *Nature* **321**, 769 (1986).
4. Guidon, N. & Delibrias, G. in *New Evidence for the Pleistocene Peopling of the Americas* (ed. Bryan, A. L.) 288–299 (Center for the Study of Early Man, Orono, 1986).
5. Dillehay, T. *Scient. Am.* **251**(4), 100–109 (1984).
6. Dillehay, T. in *New Evidence for the Pleistocene Peopling of the Americas* (ed. Bryan, A. L.) 319–337 (Center for the Study of Early Man, Orono, 1986).
7. Dillehay, T. D. et al. *J. Field Archaeol.* **9**, 547–550 (1982).
8. Pino, M. *Site and Regional Geology* (in preparation).
9. Heusser, C. J., Streeter, S. S. & Stuiver, M. *Nature* **294**, 65–67 (1981).
10. Mercer, J. H. in *Late Cretaceous Palaeo Climates of the Southern Hemisphere* (ed. Vogel, J. C.) 45–58 (SASQUA Int. Symp., Swaziland, 1983).
11. Villagran, C. *Rev. Chilena Hist. Nat.* **58**, 57–69 (1985).

Sexual differences in species recognition of avian song

W. A. Searcy & E. A. Brenowitz*

Department of Biological Sciences and Pymatuning Laboratory of Ecology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

* Department of Psychology, University of Washington, Seattle, Washington 98195, USA

Studies of species recognition in male birds have often identified features that are distinctive, invariant markers of a species song but which can be altered without altering male response^{1–4}. Such results raise the question of why natural selection has favoured the retention of these features in the song. We propose that such features are favoured by selection because females use them in species recognition even though males do not. There are excellent theoretical grounds for expecting that males and females of the same species would differ in their means of species recognition, as required by this hypothesis. We report here that female red-winged blackbirds (*Agelaius phoeniceus*) discriminate against abnormal songs previously shown to be acceptable to males of this species. Similar sexual differences in species recognition may be expected in animals other than birds, and in recognition systems based on cues other than vocalizations.

Two types of errors can be made in species recognition, which we term Type 1 and Type 2 in analogy with statistical errors. In Type 1, an individual rejects a stimulus from its own species, whereas in Type 2 a stimulus from another species is accepted. In birds, response of males to song is measured by playing songs on territories and observing the owners' aggressive reactions^{1–4}. In territory defence, males ought to be under strong selection to minimize Type-1 errors, as allowing a conspecific male to display on the territory may lead to its loss. Type-2 error would not be as important in this context, as it involves only the cost of approaching and displaying at an incorrect target. Female response to song is usually measured by playing songs to captive subjects and observing their courtship displays^{5–8}. Selection

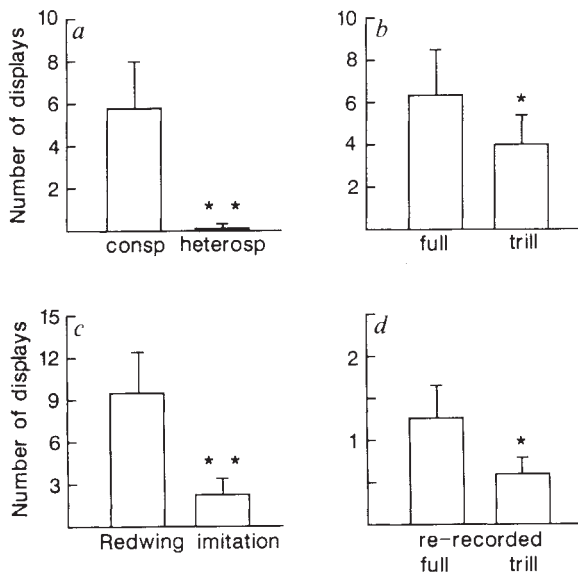


Fig. 1 Response of oestradiol-implanted female red-winged blackbirds to playback of: *a*, conspecific male song and hetero-specific male (swamp sparrow, *Melospiza georgiana*) song; *b*, full conspecific song and the trill portion only; *c*, song of a male red-winged blackbird and a mockingbird imitation of redwing song; *d*, full song and trill re-recorded at 20 m. Three minutes of song were presented at a rate of one song every 10 s in *a* and one every 15 s in *b-d*. Subjects were tested on two days with each pair of stimuli. The order of presentation was randomized on the first day and reversed on the second. Results shown are mean number of displays per female and standard error. Each subject was given an implant of silastic tubing of 1.96 mm outside diameter containing 12–14 mm of 17- β -oestradiol. Implants were placed under the skin of the back. $P < 0.05$ according to a two-tailed Wilcoxon matched pairs signed ranks test; $P < 0.01$.

should act to minimize Type-2 errors in courtship, as these could lead to wasted, hybrid matings. Type-1 errors would not be so costly, as females should have multiple opportunities to mate^{9,10}. Therefore, females should be more discriminating than males, rejecting songs that males find acceptable.

We chose to test this prediction using red-winged blackbirds, as extensive work had already been done on song recognition in males in this species. We tested captive females for response to abnormal songs previously found to be acceptable to males. Subjects were adult female red-winged blackbirds captured in Crawford County, Pennsylvania, USA during April–June, 1984–87. Our methods were similar to those used previously with sparrows^{6,7}. Females were given implants containing 17- β -oestradiol and then housed in sound attenuation chambers. Testing began one week after the oestradiol treatment. Songs were presented over speakers within the chambers, and subjects were observed through windows in the chamber doors. We counted the number of copulation solicitation displays, a courtship display in which a female red-winged blackbird crouches, raises her tail, and holds her wings out from her body. Fifteen females gave high levels of display for conspecific song and virtually none for heterospecific song (Fig. 1), confirming that responses can be interpreted as species recognition.

The song of male red-winged blackbirds consists of a few short introductory notes followed by a longer, rapid trill (Fig. 2). The response of territorial males to full song and to the trill portion only has previously been tested using a variety of aggressive behaviours to score response^{11,12}. Responses were found to be equally strong for the trill as for the full song. We presented 12 females with the same trill and full song used by Brenowitz¹² (Fig. 2). Females gave more than half again as many

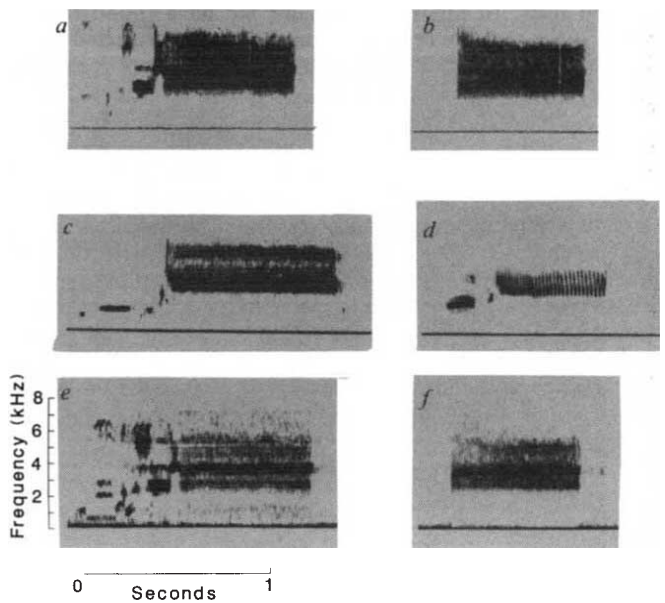


Fig. 2 Sonograms of test songs: *a*, full song of a male red-winged blackbird, recorded in Ithaca, New York, USA; *b*, the trill portion of the full song in *a*; *c*, song of a male red-winged blackbird recorded in Linesville, Pennsylvania, USA; *d*, mockingbird imitation of a male red-winged blackbird song, recorded in Reading, Pennsylvania, USA; *e*, full song in *a* re-recorded at 20 m; *f*, trill in *b* re-recorded at 20 m.

displays for the full song as for the trill (Fig. 1). The difference in response is significant according to a two-tailed Wilcoxon matched pairs test ($P < 0.05$). Thus, males should be selected to include the introductory notes to stimulate females maximally.

Brenowitz¹³ tested male red-winged blackbirds for discrimination between a mockingbird (*Mimus polyglottos*) imitation of redwing song versus a normal redwing song. The mockingbird imitation closely matches the gross temporal organization of redwing song (Fig. 2), but the fine structure of the component notes is markedly different¹⁴. Males were found to respond equally to the imitation and normal song on each of five aggressive behaviours. We tested 12 female red-winged blackbirds twice each for discrimination between a conspecific song and the same mockingbird imitation presented to males. The mean number of displays per female was four times higher for the natural song as for the imitation (Fig. 1). The difference in response was significant ($P < 0.01$). Thus, the normal fine structure of notes appears to be of benefit in stimulating females.

One problem with these results is that females were tested in sound chambers at distances of < 1 m from the speaker, whereas males were tested in noisier environments where, at least initially, they were many metres from the speaker. Thus one explanation of the results is that females were able to discriminate because they could hear the differences between the pairs of stimuli, whereas males could not discriminate because the differences between the stimuli had degraded. This explanation seems particularly likely for the trill versus full song discrimination, as the introductory notes have been shown to attenuate more rapidly with distance than the trill¹². To test this possibility, we re-recorded the full song and trill in the same type of habitat used in the trials with males, and at a distance (20 m) chosen to approximate the mean distance of males to the speaker during their playback trials (Fig. 2). We then tested 18 females for discrimination between the re-recorded stimuli. Response levels were rather low, but nevertheless females gave more than twice as many displays for the full song as for the trill ($P < 0.05$; Fig. 1), a stronger preference than shown with the original stimuli.

As predicted, female red-winged blackbirds discriminate against abnormal songs that are accepted by males in species

recognition tests. In other experiments, we have found that abnormal songs that are rejected by males are also rejected by females. We conclude that females are more discriminating than males. We cannot say whether the differences in discrimination are due simply to differing levels of motivation or whether they indicate differences in sensory and neural processing. Neurophysiological studies have suggested that sexual differences in neural pathways for song analysis may exist¹⁵. Regardless of the mechanism, the results have important implications for the interpretation of species recognition studies in general, as the theoretical arguments for expecting sexual differences apply to any species in which females respond to conspecifics in the context of courtship and males respond in the context of male-male aggression.

We thank George Reese, John Mull and Evan Ballaban for assistance. Financial support was provided by the NSF.

Received 7 December 1987; accepted 18 January 1988.

1. Falls, J. B. *Proc. XIII Int. ornith. Congr.* 259-271 (1963).
2. Emlen, S. *Behaviour* 41, 130-171 (1972).
3. Boughey, M. J. & Thompson, N. S. *Behaviour* 57, 64-90 (1975).
4. Peters, S. S., Searcy, W. A. & Marler, P. *Anim. Behav.* 28, 393-404 (1980).
5. King, A. P. & West, M. J. *Science* 195, 1002-1004 (1977).
6. Searcy, W. A. & Marler, P. *Science* 213, 926-928 (1981).
7. Searcy, W. A., Marler, P. & Peters, S. S. *Anim. Behav.* 29, 997-1003 (1981).
8. Catchpole, C. K., Dittami, J. & Leisler, B. *Nature* 312, 563-564 (1984).
9. Williams, G. C. *Adaptation and Natural Selection: A Critique of some Current Evolutionary Thought* (Princeton University Press, 1966).
10. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B. G.) 136-179 (Aldine, Chicago, 1972).
11. Beletsky, L. D., Chao, S. & Smith, D. G. *Behaviour* 73, 189-203 (1980).
12. Brenowitz, E. A. *Behav. Ecol. Sociobiol.* 10, 29-38 (1982).
13. Brenowitz, E. A. *Auk* 99, 584-586 (1982).
14. Brenowitz, E. A. *Anim. Behav.* 31, 1116-1127 (1983).
15. Williams, H. B. *Behav. neural Biol.* 44, 470-484 (1985).

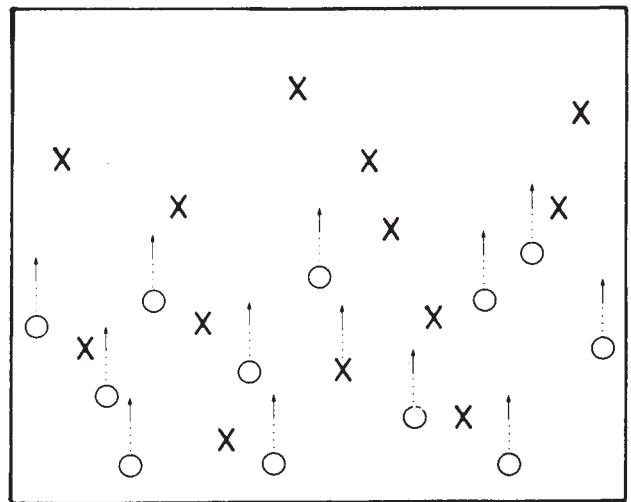


Fig. 1 The display subtended $11^\circ \times 8^\circ$. Individual letters subtended approximately $25'$. The static letters were distributed at random; the moving letters started in the bottom half of the screen and moved up. The screen was divided into 25 imaginary vertical columns, with one letter in each for the set size of 25. In smaller set sizes an adjacent subset of these columns was used to ensure the same density of letters in different conditions. A different subset was used on each trial. The display lasted 1,200 ms. Moving letters did not reach the top of the screen. The horizontal separation was sufficient to ensure that adjacent letters did not cross as the moving letters moved up the screen. There were six subjects per experimental condition. Each data point is the median of the last 150 out of a total of 200 trials per subject.

Visual search for a conjunction of movement and form is parallel

Peter McLeod, Jon Driver & Jennie Crisp

Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK

Treisman has proposed¹⁻³ when a human subject performs a visual search, the search is parallel for targets defined by a single feature, and serial for targets defined by a conjunction of features. Here we report that this is not true for targets defined by a conjunction of the features movement and form. Detection of a moving X among randomly distributed moving Os and static Xs is parallel. Search is uninfluenced by the stationary stimuli despite their spatial intermingling with the moving items. Thus, attention can be restricted to a spatially dispersed perceptual group, defined by common movement. This contradicts previous conclusions from visual search experiments^{4,5} that attention can only be assigned to contiguous regions of visual space. The search process first segregates the array into moving and stationary items, and then examines the moving group for the target form. Cells in the middle temporal region (cortical area MT) have the properties required to perform these operations^{6,7}.

In experiment 1 we used an array of moving Os and static Xs as the visual field (Fig. 1). The target (present in 50% of trials) was a single moving X. Thus only the conjunction of movement and shape distinguished the target from the non-targets. The Os and the target X (if present) moved continuously up the screen at 3.3° s^{-1} . The subjects pressed one button as soon as they detected a target; the other if they failed to find one. In different blocks of 50 trials there were 5, 9, 17 or 25 items on the screen. Figure 2 shows that the function relating detection time to set size (number of items) is essentially flat for both presence (+) and absence (○) of a target (3 ms per

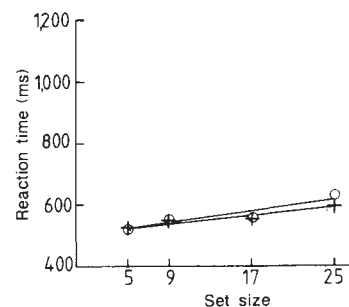


Fig. 2 Time to detect the presence or absence of a moving X in a background of moving Os and stationary Xs. Open circles, target absent; crosses, target present.

item and 5 ms per item, respectively). Thus search for a target defined by the conjunction of two of Treisman's candidate features^{1,2}, shape and movement, proceeds in parallel across the visual array.

Experiment 2 shows that a conjunction search which is well known to be slow and serial¹ can be made dramatically easier if part of the display moves. Search for an R among Ps and Qs is difficult because the target, R, shares features (either the diagonal line or the small loop) with each non-target item. It is only the unique conjunction of features which marks it as a target. In experiment 2a the Os of experiment 1 were replaced by Qs, the stationary Xs by Ps and the target X by an R. All items were stationary. Figure 3 shows the slow and serial nature of this search (dashed lines). The slopes of detection time against set size are: target present (+), 24 ms per item; target absent (○), 68 ms per item. This search task was repeated (experiment 2b) with the Qs and the R (if present) moving up the screen like the Os in experiment 1. The effect of set size on search time is greatly reduced (the solid lines in Fig. 3). The slopes of detection time against set size are: target present, 8 ms per item;

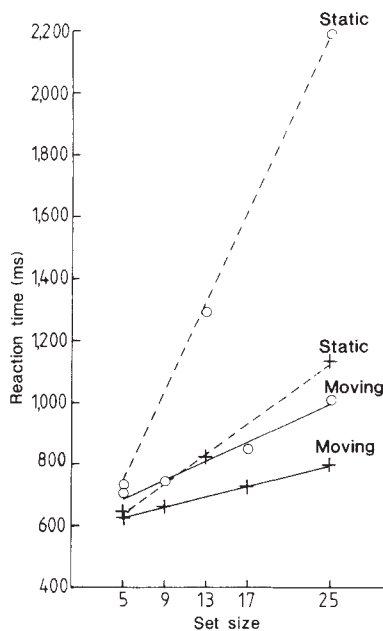


Fig. 3 Time to detect the presence (+) or absence (○) of an R in a background of Ps and Qs. Dashed lines: all items are stationary. Solid lines: Qs and the R (if present) move up the screen, the Ps are stationary.

target absent, 15 ms per item. If movement and form are independently represented features Treisman's theory does not predict this contrast between experiments 2a and 2b. The moving R in experiment 2b still has features in common with all non-target objects. Search for it should therefore be as slow as in experiment 2a.

Subjectively, the effect of moving part of the display is clear. The moving part, Os in experiment 1, Qs in experiment 2b, form a coherent group which is perceptually distinct from the background. Within this moving group the target can be distinguished from the other items by the possession of a single feature (a diagonal line in experiment 1, a small loop in experiment 2b). The parallel search we observed suggests that the visual system is successfully able to ignore the features in the static background and carry out a search limited to the moving group.

This conclusion is supported by experiment 3. The stationary Ps of experiment 2b were replaced by stationary Qs—that is, the task performed was a search for a moving R among stationary and moving Qs. According to the Treisman theory outlined above, search should be easier than it was in experiment 2b because the target (R) can be separated from all the non-targets (moving and stationary) by the possession of a unique visual feature (the small loop). But if we are correct in our conjecture that the effect of movement is to allow an initial operation which separates the visual field into two groups, moving and non-moving, followed by search within the moving group, the physical appearance of the non-moving items should have little influence on search time. Figure 4 shows that changing the static items from Ps to Qs has, as we predict, virtually no effect on the search time for the target R.

Earlier results using the conjunction search paradigm were taken as evidence for the spatial spotlight metaphor for visual attention—the idea that attention can only be directed to a contiguous region of visual space. Search for conjunctions of colour and form can be parallel but only if commonly coloured items are spatially adjacent⁴; the implication is that attention can be restricted to a spatial area but not to a particular colour. Nakayama and Silverman⁵ interpreted their observations of parallel search for conjunctions of depth with colour (for example, a near red item among near blue items and far red items), or depth with movement, as evidence for a modified

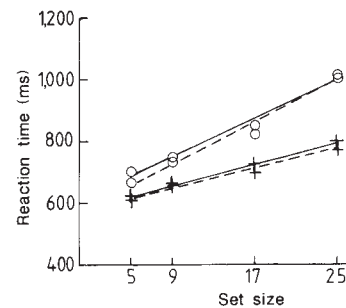


Fig. 4 Time to detect the presence (+) or absence (○) of a moving R in a background of moving Qs and stationary Ps (solid lines), or moving Qs and stationary Qs (dashed lines).

spatial spotlight view. They suggested that attention can be directed to a specific plane in three-dimensional space, so that stimuli at other depths do not influence search. They hypothesized that "the visual system can sequester processing or restrict attention in the spatial dimension but *not to other visual dimensions such as colour and motion*" (our italics). But our results show that attention can be directed to just the moving items in a spatially intermingled display of moving and stationary items. Thus, contrary to Nakayama and Silverman's hypothesis, spatial position does not have a privileged status; visual attention can be directed to perceptual groups whose components are spatially dispersed.

We have shown here that the visual system can segregate randomly distributed displays of moving and stationary stimuli into two groups; the moving and the stationary. It can then conduct a search for the presence of a form feature among the moving stimuli without influence from the features present in the stationary group. The properties of cells in cortical area MT suggest that they could carry out such a process. Cells in MT are responsive to moving stimuli but show little response to stationary ones⁶. Thus they act as a 'movement filter'. All moving stimuli, irrespective of their position on the display, are represented there; stationary stimuli are represented weakly, if at all. Many of these cells are also orientation sensitive⁷. Thus the presence of a moving X among stationary Xs and moving Os can be signalled by activity of cells in MT sensitive to a line at 45°. If such cells become active, a yes response is required. The latency of this response will not depend either on set size or the features present in the stationary group.

The proposal of a system which represents moving stimuli but is insensitive to stationary stimuli is consistent with Tolhurst's⁸ finding that exposure to stationary sinusoidal gratings has little effect on the threshold for subsequent moving gratings. In contrast, he found that exposure to moving gratings affects thresholds for moving and stationary gratings equivalently. This implies that the system analysing stationary forms is influenced by moving forms. We confirm this asymmetry in a variation of Experiment 1. The subject's task was now to search for a stationary X among stationary Os and moving Xs. This proved more difficult than search for a moving X among moving Os and stationary Xs had been. The slope for positive responses was 7 ms per item, and for negatives 26 ms per item (compared to 3 and 5 ms per item in experiment 1).

J.C. was supported by a grant from the Oxford University Research and Equipment Committee.

Received 18 January; accepted 5 February 1988.

1. Treisman, A. & Gelade, G. *Cognitive Psychol.* **12**, 97-136 (1980).
2. Treisman, A. *Sci. Amer.* **255**(5), 106-115 (1986).
3. Treisman, A. in *Handbook of Perception and Human Performance* Vol. 2 (eds Boff, K., Kaufman, L. & Thomas, J.) Ch. 35 (Wiley, New York, 1986).
4. Treisman, A. *J. exp. Psychol.: Hum. Percept. Perf.* **8**, 194-214 (1982).
5. Nakayama, K. & Silverman, G. *Nature* **320**, 264-265 (1986).
6. Maunsell, J. & Van Essen, D. *J. Neurophysiol.* **49**, 1127-1147 (1983).
7. Livingstone, M. & Hubel, D. *J. Neurosci.* **7**(11), 3416-3468 (1987).
8. Tolhurst, D. *J. Physiol., Lond.* **231**, 385-402 (1973).

A physiological role for GABA_B receptors in the central nervous system

P. Dutar* & R. A. Nicoll†

Departments of Pharmacology and Physiology,
University of California, San Francisco, California 94143, USA

The role of GABA in synaptic transmission in the mammalian central nervous system is more firmly established than for any other neurotransmitter. With virtually every neuron studied, the synaptic action of GABA is mediated by bicuculline-sensitive GABA_A receptors which selectively increase chloride conductance. However, it has been shown that GABA has a presynaptic inhibitory action on transmitter release that is insensitive to bicuculline and is selectively mimicked by baclofen¹. The receptors involved in this action are referred to as GABA_B receptors, to distinguish them from the classic bicuculline-sensitive GABA_A receptors. In hippocampal pyramidal cells an additional postsynaptic action of GABA and baclofen has been reported that is also insensitive to GABA_A antagonists^{2,3}, and may be mediated by GABA_B receptors on the postsynaptic neuron. This action of GABA and baclofen involves an increase in potassium conductance²⁻⁷. Synaptic activation of pathways converging on hippocampal pyramidal cells results in a slow inhibitory postsynaptic potential which involves an increase in potassium conductance^{4,8,9}, and it has been suggested that GABA_B receptors might be responsible for this synaptic potential⁴. However, to establish convincingly that GABA_B receptors are physiologically important in the central nervous system, a selective GABA_B antagonist is required. Here we provide this missing evidence. Using the hippocampal slice preparation, we now report that the phosphonic acid derivative of baclofen, phaclofen¹⁰, is a remarkably selective antagonist of both the postsynaptic action of baclofen and the bicuculline-resistant action of GABA, and that it selectively abolishes the slow inhibitory postsynaptic potential in pyramidal cells.

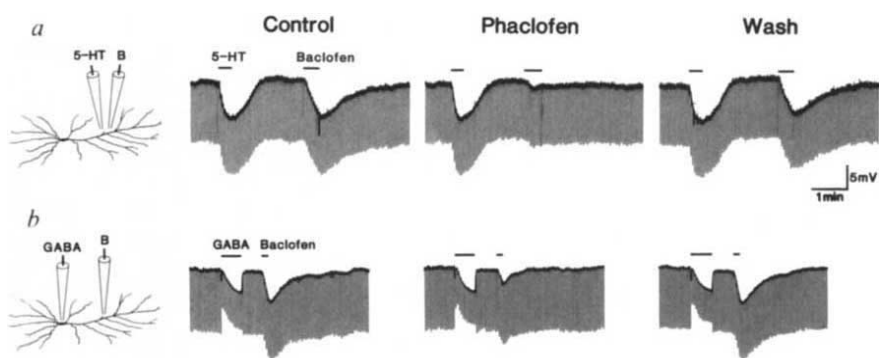
* Present address: Laboratoire de Neurophysiologie Pharmacologique, INSERM U161, 2 rue d'Alesia, 75014 Paris, France.

† To whom correspondence should be addressed.

The first series of experiments was carried out to determine if phaclofen could antagonize the postsynaptic action of baclofen in CA₁ hippocampal pyramidal cells. Phaclofen (0.2–1 mM) has no consistent effect on the resting membrane potential, input resistance, the shape of the action potential or the afterhyperpolarization which follows a series of action potentials. As relatively high concentrations of the drug are required for its antagonistic action¹⁰, it is essential to establish that its action is selective. We therefore compared the effect of phaclofen in the same cell on GABA_A responses and on 5-HT (5-hydroxytryptamine) responses. 5-HT responses were chosen because 5-HT_{1A} receptors activate the same potassium conductance as GABA_B receptors¹¹. Phaclofen (0.3–0.5 mM) antagonized the response to ionophoretic application of baclofen by 70% (±6) (*n* = 10), but has no effect on the response to ionophoretic application of 5-HT (*n* = 7) to the same dendritic site (Fig. 1a). The antagonism occurs rapidly (5–10 min), and recovers quickly (5–15 min) after washing the slice in a drug-free medium. GABA_A responses on ionophoretic application of GABA to soma⁴ are also unaffected by phaclofen (*n* = 2) (Fig. 1b).

We next compared the properties of the baclofen response to those of the bicuculline-resistant GABA response, both of which result from an increase in potassium conductance²⁻⁷. To determine if the potassium conductance activated by baclofen is the same as that activated by GABA we have examined whether the currents underlying these responses are non-additive. If the increase in potassium conductance evoked by these two agonists results from entirely distinct mechanisms, the maximal activation of the current by one agonist would not be expected to affect the current evoked by the other agonist. As shown in Fig. 2a, this is not the case. In the presence of GABA_A antagonists, the outward current evoked by brief ionophoretic applications of GABA to the dendrites is essentially completely occluded by bath application of a high concentration of baclofen. This result suggests that both agonists activate the same potassium conductance. Ogata *et al.*¹³ have suggested that the two conductances might be different. Since in their studies the baclofen response, but not the GABA response, showed inward rectification and was blocked by low concentrations 4-aminopyridine (4-AP). Other studies have reported inward rectification for the GABA response^{4,5} and we (unpublished observations) do not find the

Fig. 1 Phaclofen selectively antagonizes the baclofen-induced hyperpolarization. **a** (left), Serotonin (5-HT) and baclofen (B) are ionophoresed onto the dendrites of the pyramidal cell from electrodes positioned at the surface of stratum radiatum, 200 µm away from the cell layer. On the right, intracellular recording shows the hyperpolarizing responses to 5-HT (110 nA) and baclofen (120 nA) applications. In the presence of phaclofen (0.5 mM), superfused for 6 min, the 5-HT response remains unchanged; in contrast, the baclofen response is blocked. After washing out the phaclofen (6 min), the baclofen response recovers. Downward deflections represent constant current hyperpolarizing pulses applied throughout. **b**, GABA electrode positioned close to the cell body to elicit a response mediated by GABA_A receptors⁴. The baclofen electrode (B) was positioned on the surface of the stratum radiatum. Intracellular recording on the right shows that somatic application of GABA (20 nA) and dendritic application of baclofen (110 nA) induce hyperpolarizing responses. In the presence of phaclofen (0.7 mM) superfused for 7 min, the somatic GABA response is unchanged and the baclofen response is strongly depressed. This response recovers 7 min after washing out the phaclofen. **a** and **b** show recordings from two different cells. Resting membrane potential –58 mV (**a**) and –61 mV (**b**). The drugs were ionophoresed for the periods indicated by the bars. Calibration in **a** applies also to **b**.



Methods. Experiments were performed on rat hippocampal slices prepared by standard methods¹². Briefly, intracellular current-clamp and single-electrode voltage-clamp recordings were made from CA₁ pyramidal cells using an AxoClamp II amplifier. For current-clamp experiments, microelectrodes were filled with potassium methylsulphate (2 M; 80–120 MΩ). For voltage-clamp experiments, electrodes were filled with KCl (3 M; 20–80 MΩ) and tetrodotoxin (0.3 µM) was added to the perfusing medium to block action potentials and synaptic potentials. The ionophoretic electrodes were filled with either GABA (1 M, pH 5, Sigma), (±) baclofen (40 mM, pH 3, Ciba Geigy) or 5-HT (creatinine sulphate) (40 mM, pH 4, Sigma) dissolved in distilled water. GABA_A antagonists bicuculline methiodide (40–100 µM, Sigma) and picrotoxin (10–40 µM, Sigma), as well as phaclofen (3-amino-2-(4-chlorophenyl)-propyl phosphonic acid) (0.2–1.5 mM, Tocris Neuramin), were dissolved in the medium just before use. Synaptic responses, including the fast and the slow IPSP were evoked by electrical stimulation (0.1 ms pulse width) of Schaffer collateral/commissural fibres using bipolar stimulating electrodes located in stratum radiatum.

Fig. 2 Characterization of bicuculline-resistant GABA response. *a*, A pyramidal cell was voltage-clamped at -60 mV. Pulses of GABA (60 nA) were applied by ionophoresis in the dendritic field in the presence of bicuculline methiodide (40 μ M) and picrotoxin (20 μ M). Outward currents resulting from GABA application were recorded. Baclofen (100 μ M), added to the bath during the time indicated by the bar, induces an outward current. Applications of GABA during the baclofen response evoke a minimal response, indicating a non-additivity of the currents. *b* and *c*, Schematic location of ionophoretic electrodes is shown on the left. To the right, GABA ionophoresed in stratum radiatum in the presence of GABA_A antagonists bicuculline methiodide (BMI) (40 μ M) and picrotoxin (Picro) (20 μ M) evokes hyperpolarizing responses (control traces). Phaclofen (0.5 mM) applied for 15 min in the bath antagonizes both the baclofen response (*b*) and the bicuculline-resistant GABA responses (*b*, *c*). In contrast, the 5-HT response is not affected (*c*). Twenty minutes after washing the phaclofen from the bath, the baclofen and GABA responses recovered. *b* and *c*, Two different cells are represented; *b*, resting membrane potential, -63 mV, GABA 80 nA, baclofen 140 nA; *c*, resting membrane potential -61 mV, GABA 120 nA, 5HT 100 nA. Drugs were ionophoresed during the times indicated by the bar. Calibrations in *b* apply also to *c*.

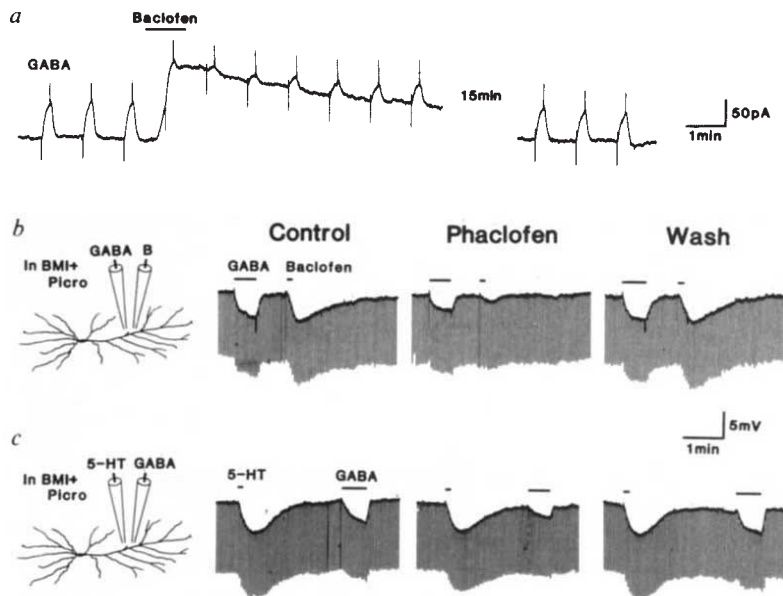
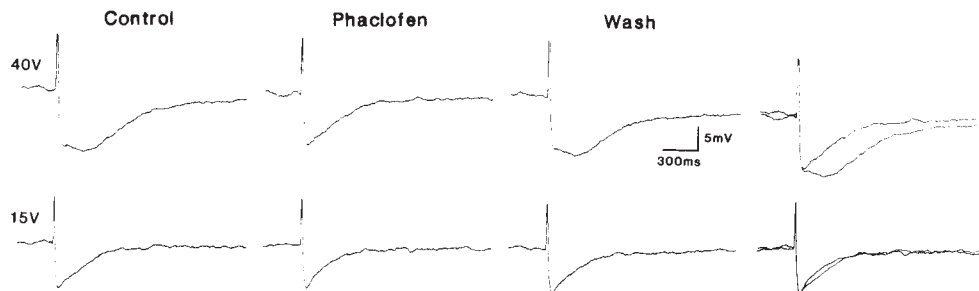


Fig. 3 Phaclofen selectively blocks the slow i.p.s.p. Synaptically evoked potentials induced by orthodromic stimulation in stratum radiatum are recorded from a CA₁ pyramidal cell. In the top trace, the stimulation intensity (40 V) evokes a biphasic hyperpolarizing response which includes both a fast and a slow i.p.s.p. At lower intensity (15 V), only the fast i.p.s.p. is elicited (bottom trace). In the presence of phaclofen (0.2 mM), applied to the bath for 5 min, the slow i.p.s.p. is abolished (top). In contrast, the fast i.p.s.p. is unchanged (bottom). Top and bottom traces are recordings from the same cell. Resting membrane potential, -56 mV. The records on the right are superimposed tracings of the control response and the response in phaclofen.



high degree of sensitivity of the baclofen response to 4-AP. To determine if the two agonists also share the same receptor we have examined the effect of phaclofen on the bicuculline-resistant GABA response. Phaclofen, at the same concentrations that antagonize the baclofen response, also antagonize the GABA response evoked in the presence of GABA_A antagonists by 36% (± 15) ($n = 13$) (Fig. 2*b*). This antagonism is also selective, because the response to 5-HT, compared in the same cell, is unaffected (Fig. 2*b*). These data suggest that GABA and baclofen act at the same receptor.

For a given concentration of phaclofen, the baclofen response was more sensitive than the bicuculline-resistant GABA response. One possible explanation for this difference is that uptake mechanisms exist for GABA but not for baclofen. Therefore to produce an ionophoretic response of a given size, GABA will act on a smaller area of membrane at higher concentrations than baclofen. A competitive antagonist such as phaclofen will be less effective against responses generated from the peak of the dose-response curve.

The results presented thus far indicate that GABA and the GABA_B agonist, baclofen, activate the same potassium conductance in hippocampal pyramidal cells and that responses to both agonists are selectively reduced by the antagonist phaclofen. It was therefore of interest to consider whether phaclofen might antagonize the slow inhibitory postsynaptic potential (i.p.s.p.) recorded in hippocampal pyramidal cells, as this i.p.s.p. is resistant to GABA_A antagonists and is generated by an increase in potassium conductance. Indeed, phaclofen (0.2–0.5 mM) dramatically reduced the slow i.p.s.p. in every experiment, but had little or no effect on the fast i.p.s.p. which is mediated by GABA_A receptors ($n = 8$). Figure 3 shows results from an experiment in which an afferent input was stimulated at two different

intensities. At the high intensity (40 V), a fast and slow i.p.s.p. are generated, and phaclofen selectively blocks the slow component. At the low intensity (15 V), the response consists primarily of the fast, GABA_A, i.p.s.p. and phaclofen has little effect on this response, indicating that phaclofen has not decreased the release of GABA onto the postsynaptic neuron. This experiment provides strong support for the hypothesis that the slow i.p.s.p. is mediated by the activation of GABA_B receptors on the postsynaptic membrane.

In summary, although GABA_B receptors have been known for some time to exist in the CNS, the present experiments establish for the first time a physiological role for GABA_B receptors. Specifically, GABA released from interneurons interacts with two subtypes of GABA receptors on the membrane of pyramidal cells: the GABA_A receptor which generates the fast i.p.s.p., and the GABA_B receptor which generates the slow i.p.s.p. The synaptic activation of these two inhibitory receptors depends on the strength of the afferent input. Weak afferent inputs evoke only the fast i.p.s.p., presumably because the interneurons do not fire a sufficient number of action potentials to release enough GABA to activate the GABA_B receptors, which require higher concentrations of GABA for their activation⁴. GABA_B receptors would only be activated with strong afferent inputs. Slow i.p.s.p.s, mediated by increase in potassium conductance^{14–20}, as well as GABA_B binding²¹, have been described throughout the CNS, and it will be of interest to determine if these potentials are also mediated by GABA_B receptors.

We thank Drs D. I. B. Kerr and J. Ong for the generous sample of phaclofen used in the initial studies and Drs J. Kauer and R. Malenka for comments on the manuscript. This work was supported by the NIH (R.A.N.) and the Fogarty International Center for NIH and the IPSEN Foundation (P.D.).

Received 2 November 1987; accepted 7 January 1988.

- Hill, D. R. & Bowery, N. G. *Nature* **290**, 149–152 (1981).
- Newberry, N. R. & Nicoll, R. A. *Nature* **308**, 450–452 (1984).
- Blaxter, T. J. & Cottrell, G. A. Q. *Jl exp. Physiol.* **70**, 75–93 (1985).
- Newberry, N. R. & Nicoll, R. A. *J. Physiol., Lond.* **360**, 161–185 (1985).
- Gahwiler, B. H. & Brown, D. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1558–1562 (1985).
- Inoue, M., Matsuo, T. & Ogata, N. *Br. J. Pharmac.* **84**, 833–841 (1985).
- Inoue, M., Matsuo, T. & Ogata, N. *Br. J. Pharmac.* **86**, 515–524 (1985).
- Alger, B. E. *J. Neurophysiol.* **52**, 892–910 (1984).
- Hablitz, J. J. & Thalmann, R. H. *J. Neurophysiol.* **58**, 160–179 (1987).
- Kerr, D. J. B., Ong, J., Prager, R. H., Gynther, B. D. & Curtis, D. R. *Brain Res.* **405**, 150–154 (1987).
- Andrade, R., Malenka, R. C. & Nicoll, A. *Science* **234**, 1261–1265 (1986).
- Nicoll, R. A. & Alger, B. E. *J. Neurosci. Meth.* **4**, 153–156 (1981).
- Ogata, N., Inoue, M. & Matsuo, T. *Synapse* **1**, 62–69 (1987).
- Connors, B. W., Gutnick, M. J. & Prince, D. A. *J. Neurophysiol.* **48**, 1302–1320 (1982).
- Satou, M., Mori, K., Tazawa, Y. & Takagi, S. *J. Neurophys.* **48**, 1142–1156 (1982).
- Mori, K., Nowicky, M. C. & Shephard, G. M. *J. Physiol., Lond.* **314**, 295–309 (1981).
- Jahr, C. E. & Nicoll, R. A. *J. Physiol.* **326**, 213–234 (1982).
- Howe, J. R., Sutor, B. & Zieglansberger, W. *J. Physiol., Lond.* **384**, 539–569 (1987).
- Avoli, M. *Brain Res.* **370**, 165–170 (1986).
- Stevens, D. R., Gallagher, J. P. & Shinnick-Gallagher, P. *Synapse* **1**, 184–190 (1987).
- Bowery, N. G., Hudson, A. L. & Price, G. N. *Neuroscience* **20**, 365–383 (1987).

A synthetic vaccine protects humans against challenge with asexual blood stages of *Plasmodium falciparum* malaria

Manuel E. Patarroyo*, Roberto Amador*, Pedro Clavijo*, Alberto Moreno*, Fanny Guzman*, Pedro Romero*, Ricardo Tascon*, Antonio Franco*, Luis A. Murillo*, Gabriel Ponton† & Gustavo Trujillo†

* Instituto de Immunologia, Hospital San Juan de Dios, Universidad Nacional de Colombia, Bogota, Colombia SA.
† Hospital Militar Central de Colombia, Fuerzas Militares de Colombia, Bogota, Colombia S.A.

We have previously shown that a mixture of three synthetic peptides (83.1, 55.1, 35.1), corresponding to fragments of the relative molecular mass 83,000 (83K), 55K and 35K *Plasmodium falciparum* merozoite-specific proteins, induces protection in *Aotus trivirgatus* monkeys experimentally infected with *P. falciparum*^{1,2}. Here we describe two polymeric synthetic hybrid proteins based on these peptides that delay or suppress the development of parasitaemia in immunized human volunteers.

In the search for a vaccine against the different infectious stages of malaria caused by *P. falciparum*, we designed two hybrid protein polymers (Fig. 1). These contain epitopes known to induce partial or complete protection in experimentally-infected *A. trivirgatus* monkeys or that were good candidates for a malaria vaccine^{1–5}. These molecules were pure as observed by HPLC (Fig. 2a) and were caused to polymerize into synthetic proteins by the addition of cysteines at the N and C-terminals and glycine and proline as spacers. The relative molecular mass (M_r) of SPf(66)30 was 150K and of SPf(105)20 was 100K, indicating that the polymers were of 30 and 20 units, respectively (Fig. 2b).

The safety, immunogenicity and protectivity of SPf(66)30 as well as its individual peptides was previously assayed in monkeys and no histopathological, behavioural or clinical laboratory abnormalities were observed in any of the immunized and killed animals. The immunogenicity and protectivity of some of the individual peptides of SPf(105)20 had been previously assayed by us¹ and by others⁵ in monkeys. To examine the safety and immunogenicity of these synthetic proteins in man and the protection afforded against malaria, 13 male volunteers (aged 18–21 years) were selected from 109 healthy, high-school graduate volunteer soldiers from the Colombian Military Forces. On the basis of their clinical history, origin from areas where malaria is not endemic, clinical status and laboratory tests (haematocrit, total and differential blood-cell count, complete serum chemistry, urinalysis and serological tests for the

Table 1 Humoral and cellular immune responses in vaccinees

Volunteer	Immunizing antigen	Antibody titres assessed by ELISA				IIFA	
		SPf(66)30 P.I.	SPf(66)30 P.C.	SPf(105)20 P.I.	SPf(105)20 P.C.	Schizonts P.I.	Schizonts P.C.
W.B.	SPf(66)30	0.140	0.580	0.160	0.300	0	40
W.G.	SPf(66)30	0.260	0.570	0.310	0.545	0	40
L.C.	SPf(66)30	0.340	0.640	0.160	0.210	20	160
D.A.	SPf(66)30	0.08	0.750	0.100	0.320	0	40
J.C.	SPf(66)30	0.160	0.170	0.110	0.170	0	20
H.A.	SPf(105)20	0.240	0.270	0.300	1.000	20	40
J.S.	SPf(105)20	0.170	0.480	0.110	0.440	0	160
E.O.	SPf(105)20	0.110	0.120	0.240	0.720	0	160
H.B.	SPf(105)20	0.180	0.570	0.00	1.200	20	40
A.C.	Saline	0.240	0.280	0.210	0.190	0	0
J.D.	Saline	0	0	0.007	00	0	0
C.B.	Saline	0.200	0.270	0.160	0.440	0	0
J.E.	—	ND	ND	ND	ND	0	0

Volunteer	Immunizing antigen	PBMC proliferation assays				Schizonts	
		SPf(66)30 P.I.	SPf(66)30 P.C.	SPf(105)20 P.I.	SPf(105)20 P.C.	P.I.	P.C.
W.B.	SPf(66)30	0.6	3.1	ND	3.1	2.0	1.5
W.G.	SPf(66)30	1.7	5.6	ND	4.1	1.7	5.6
L.C.	SPf(66)30	0.6	5.5	ND	3.0	1.4	12.4
D.A.	SPf(66)30	1.1	13.9	ND	7.8	1.1	6.6
J.C.	SPf(66)30	1.0	16.1	ND	9.5	1.1	9.1
H.A.	SPf(105)20	0.4	18.9	ND	18.3	0.7	35.1
J.S.	SPf(105)20	1.3	12.0	ND	5.3	1.1	13.4
E.O.	SPf(105)20	1.7	0.6	ND	0.6	3.0	0.8
H.B.	SPf(105)20	ND	1.8	ND	2.3	ND	3.3
A.C.	Saline	0.1	1.3	ND	1.3	0.2	1.4
J.D.	Saline	2.1	1.7	ND	0.8	2.5	1.0
C.B.	Saline	1.3	3.5	ND	3.4	1.8	4.0
J.E.	—	ND	ND	ND	ND	ND	ND

Fluorescent antibody titres are expressed as the reciprocal of the final serum dilution. Air-dried *P. falciparum* infected erythrocyte cultures with 10% parasitaemia at the schizont-merozoite stages were used for schizont-IIFA¹⁵ and human glutaraldehyde-fixed *P. falciparum* ring-infected erythrocytes at 12% parasitaemia for Pf155/RESA-IIFA¹⁵. Antibodies against RESA were negative with all sera. Fluorescence-conjugated F(ab)₂ fragments of affinity-purified goat anti-human IgG were used. Cooke ELISA plates (Dynatech Corp.) were coated with 10 µg per well of each of the molecules dissolved in carbonate coating buffer, pH 9.6. Plates were blocked for two hours with phosphate buffer solution (PBS) pH 7.3 with casein and washed 3 times with PBS containing 0.05% Tween 20. The plates were incubated overnight at 37 °C in serum diluted in 1:50 in PBS + 1% casein and washed 3 times as before. Peroxidase-conjugated goat anti-human IgG was added at a dilution of 1:1,000 in PBS, incubated for 1 h, washed 3 times and the peroxidase substrate *o*-phenylenediamine was added. Optical density was read at 490 nm (ref. 16) and considered positive if above 0.200. All sera were anti-NANP negative, perhaps because NANP was included in the synthetic hybrid proteins once instead of the 3× tandem repeat. This experiment was carried out in Dr. R. Nussenzweig's laboratory by one of us (P.R.). Proliferation of human peripheral blood mononuclear cells (PBMC) was assayed using PBMCs isolated by Ficoll-Hypaque centrifugation and seeded at 150,000 and 300,000 cells in 0.2 ml RPMI + 10% pooled 0+ human serum, 25 mM HEPES, penicillin, streptomycin and either the protein polymer or schizont sonicates at concentrations of 1, 5 and 12.5 µg ml⁻¹. On the fifth day of *in vitro* culture, 0.5 µCi of (³H)thymidine (Amersham, 93 Ci mmol⁻¹) was added to each well. The cells were harvested 16 h later and incorporated ³H was assessed by β-emission spectroscopy^{17,18}. Only stimulation indexes above 5.0 were considered positive. The FVO and several Colombian wild strains provided the material for the humoral and cellular immune test. P.I., pre-immune; P.C., pre-challenge.

presence of antibodies to hepatitis B virus and *P. falciparum*), 13 volunteers, who chose to cooperate in this study were accepted. No incentives such as money or promotions were offered. All had normal laboratory tests and were in excellent mental and physical condition. Based on WHO recommendations for human trials^{6,7}, written consent, with the understanding that

Table 2 Development of postchallenge parasitaemia in vaccinees

	Percentage of parasitaemia on days after challenge													
	5		6		7		8		9		10		11	
	a.m.	p.m.	a.m.	p.m.	a.m.	p.m.	a.m.	p.m.	a.m.	p.m.	a.m.	p.m.	a.m.	p.m.
SPF(66)30														
W.B.	0	0	0	0	0.002	0.007	0	0	0	0	0	0	0	0
W.G.	0.015	0.036	0.040	0.002	0.440	0.289	0.020	0	0.171	0.368	0.019	0.006	0.190	0.120
L.C.	0	0.034	0.008	0	0.052	0.142	0.030	0.002	0.214	0.465	0.011	0.006	0.051	0.175
D.A.	0.006	0.010	0.028	0.011	0.410	0.400	0.066	0.007	0.288	0	0.014 ^T			
J.C.	0	0.006	0.007	0	0.002	0.171	0.183	0.002	0.078	2.150 ^T				
SPf(105)20														
H.A.	0	0.005	0.02	0	0.016	0.103	0.118	0.004	0.330	0.236	0.092	0.027	0.022	0.660
J.S.	0.010	0.180	0.025	0	0.002	0.220	0.013	0.005	0.603	0.660	0.011	0.02	0.384	0.420
E.O.	0	0.006	0.002	0.004	0.015	0.112	0.100	0.005	1.200 ^T					
H.B.	0.002	0.100	0.054	0.002	0.110	0.400	0.460	0.008	2.120 ^T					
Controls														
A.C.	0	0.032	0.022	0.007	0.016	0.270	0.150	0	0.115	1.600 ^T				
J.D.	0	0.033	0.026	0.005	0.008	0.870	0.780	0.017	4.260 ^T					
C.B.	0	0.120	0.072	0	0.300	2.300 ^T								
J.E.	0.008	0.100	0.115	0.010	1.700	3.600 ^{T*}								
	12		13		14		15	16	17	18	19	20	21	22
	a.m.	p.m.	a.m.	p.m.	a.m.	p.m.								
SPF(66)30														
W.B.	0	0	0	0	0	0	0	0	0	0	0	0	0	0
W.G.	0.015	0.166	0.292	0.094	0.075	0.072	0.001	0	0.020	0.030	0	0.020	0	0
L.C.	0.130	0.013	0.120	0.242	0.060	ND	0.02	0.116	0.064	0	0	0	0	0
SPf(105)20														
H.A.	0.110	0.180	0.486	0.500	0.354	0.675 ^T								
J.S.	0.06	0.190	1.050 ^T											

Group 2 (W.B., W.G. and L.C.) received the second of the 2 doses of SPf(66)30 80 days before challenge. Group 1 (D.A. and J.C.) received the last of three injections of SPf(66)30 60 days before challenge, as did the individuals vaccinated with SPf(105)20 or saline solution. On the day of challenge volunteers were intravenously inoculated with a wild *P. falciparum* strain, grade I chloroquine resistant and completely sensitive to sulphadoxine and pyrimethamine, like most Colombian wild strains. The infected red blood cells were obtained from a malaria-naïve volunteer (EG) previously inoculated with the thawed strain, blood group 0+, compatible with all the recipients and in excellent medical, clinical and laboratory condition, with all serological tests negative. The inoculum given to the volunteers was one million fresh live ring-infected erythrocytes diluted in ~4 ml sterile saline solution. Parasitaemia was monitored after the third day by thick and thin blood smears stained with Giemsa, Field and Acridine Orange every 12 h, and communicated to the volunteers by the medical staff, to allow withdrawal. Volunteers with parasitaemias above 0.5% received immediate treatment with chloroquine followed by sulphadoxine and pyrimethamine. Due to their excellent clinical condition, three volunteers (H.A., J.S. and J.D.) with parasitaemia slightly above 0.5% were left untreated for a further 24–80 h. During this time they showed partial control of parasitaemia. The individual parasitaemia evolution is shown for better understanding of the protective effect of the synthetic hybrid proteins. During the first four days all parasitaemias were negative. Giemsa staining readings were lower than Acridine Orange. (Data not shown.) T, infection terminated. * This sample was taken just before and 2 h after treatment began.

volunteers were free to withdraw at any point of the investigation, was obtained after explaining in detail the nature of the study and its potential risk and benefits. Throughout the study volunteers were questioned periodically about their willingness to continue. This study was approved by medical committees of the Colombian Military Forces and the Colombian Ministry of Public Health. For challenge studies, volunteers were hospitalized at the Central Military Hospital of Colombia in Bogota, with full access to all medical services and intensive care (R.A. is a critical care specialist). In addition, during the hospitalization period, six physicians, authors of this paper, were on 24-hour call at the hospital. Blood smears from the subjects were analysed independently three times: by us; by the scientific staff of the Central Military Hospital of Colombia; and by the Colombian Malaria Eradication Service (SEM). We imposed on ourselves the criterion that vaccinees who developed parasitaemias greater than 0.5% would be treated with chloroquine followed with sulphadoxine and pyrimethamine or quinine.

The volunteers were divided into five groups: (1) two (D.A. and J.C.) received three 2-mg doses of SPf(66)30 on days 0, 60, and 80; (2) three (W.B., W.G. and L.C.) received the same protein and dose on days 0 and 60 only; (3) four (H.A., J.S., E.O. and H.B.) received three 2-mg doses of SPf(105)20 on days 0, 20 and 45; (4) three controls (A.C., J.D. and C.B.) received saline solution on days 0, 20 and 45; (5) one volunteer (J.E.) served as naïve receptor for passing on the *P. falciparum* strain.

Both synthetic hybrid proteins and the saline solution were adsorbed to Al(OH)₃ before inoculation.

One subject vaccinated with SPf(66)30 (D.A.) developed a general urticarial eruption five minutes after the third immunization. No hypotension or dyspnoea developed and the rash responded rapidly to treatment with hydrocortisone and adrenalin. The cause of this reaction is unknown but was attributed to problems during the dialysis of a batch of SPf(66)30 to remove salts, Tris and dithiothreitol (DTT). As we could not detect IoE specific antibodies against the construct, as assayed by dot-blot-ELISA, this dialysis was therefore omitted from the purification protocol and the desalting performed by gel filtration.

Neither the subject vaccinated three times with another batch of SPf(66)30 nor the ones vaccinated three times with SPf(105)20 developed severe, systemic or local side effects. Slight pain, local erythema and induration at the inoculation site were, however, noted in all the subjects. None of the volunteers presented fever or showed significant changes in blood-cell count, blood chemistry and urinalysis on days -1, 1, 3 and 5 after each immunization. Autoimmunity tests (rheumatoid factor, antinuclear antibodies, Coombs test and antimyocardial fibre antibodies) were systematically negative.

To study the dynamics of the humoral and cellular immune responses, blood samples were taken the day before each immunization, 15 days later, and the day before the challenge.

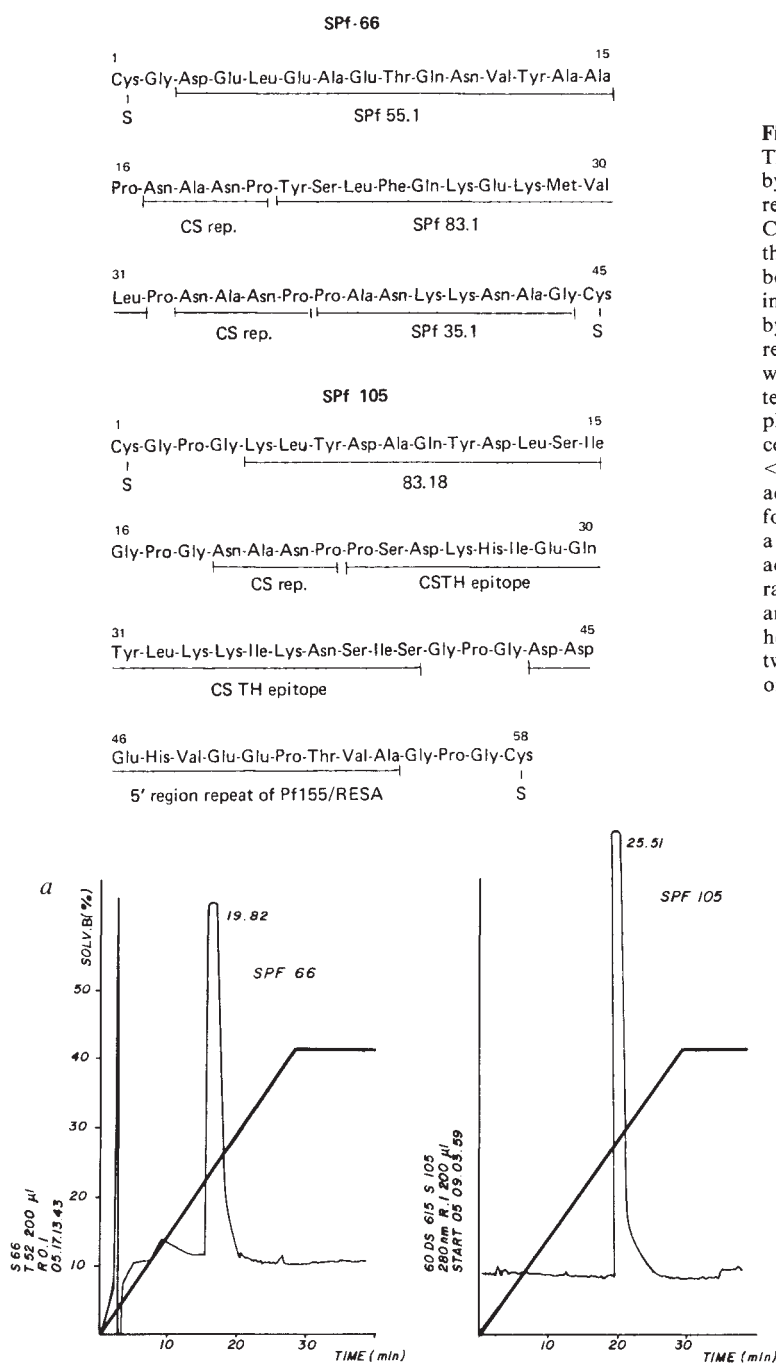


Fig. 2 *a*, HPLC chromatograms of the synthetic malaria proteins. Solvent A, 5%, CH₃CN, 0.068% trifluoroacetic acid (TFA) in deionized water; solvent B, 90% CH₃CN, 0.045% TFA. Columns used were octadecylsilane (ODS) at 45 °C. *b*, 15–5% gradient SDS-PAGE of SPf(66)30, showing the degree of polymerization of this protein as an example.

In this last sample, antibody titres were determined by peptide-antipeptide enzyme-linked immunosorbent assay (ELISA) using the synthetic protein molecules as antigens. No antibodies were detected against the circumsporozoite repeat molecule (NANP) in any sera or against Pf155/RESA (rine erythrocyte surface antigen)^{5,15} in sera from individuals immunized with SPf(105)20 (Table 1). Indirect immunofluorescence assays (IIFA) showed that all sera contained antibodies to merozoite-schizonts in titres between 1:20 and 1:160. All preimmune sera and sera from the controls and naive receptors were negative or had antibody titres below 1:20. No correlation was found between antibody levels and anti-malarial protection, but the only person vaccinated with SPf(66)30 who was not protected (J.C.) had the lowest antibody titres by all methods. Proliferation assays of peripheral blood mononuclear cells (PBMC) using either protein hybrid

Fig. 1 Amino-acid sequence of the synthetic hybrid molecules. The molecules used for immunization were chemically synthesized by the Merrifield method¹² in a Beckman 990 synthesizer. The resin used was *p*-methyl-benzhydrylamine HCl (US Biochemical Corp.). The efficiency of each coupling reaction was assessed using the ninhydrine test¹³. Reactions were repeated if efficiencies fell below 99%. In SPf(66)30, peptide 35.1 is three amino acids shorter in its N-terminal portion. Peptides were liberated from the resin by treatment with HF Low-High¹⁴ for 2 h at 0 °C and 1 h at –20 °C, respectively; using *p*-cresol as scavenger. The final product was washed 10 times with 10 ml of ethyl-ether and the peptides extracted with 10% acetic acid. After evaporating the acid, peptides were placed in reducing conditions, purified by HPLC on reverse phase columns (ODS) and freeze-dried. The peptides, which had *M_s* of <14–15K, were solubilized in 10 mM Tris-HCl pH 7.6. DTT was added in equimolar quantities to reduce the disulphide bonds formed during freeze-drying. The peptide solution was placed in a water bath at 60 °C for 15 min. The pH was adjusted to 4.0 by addition of HCl during the last two minutes. The solution was rapidly cooled and the pH modified to 7.6 with 0.1M Tris-HCl and then vigorously oxygenated blowing pure oxygen for 6–10 hours. The peptides were either dialysed under sterile conditions, twice against 1 mM Tris-HCl and 6 more times in distilled water, or desalted with Sephadex G-25. They were then freeze-dried and analysed using SDS-PAGE (see Fig. 2*b*).

SPf(66)30 or sonicates of purified schizonts as antigens showed stimulation indexes (SI) below 3.0 before the first vaccination. After each vaccination and before the challenge, the stimulation indexes varied from 0.61 to 35.1 but did not correlate with either antibody titres or anti-malarial protection (Table 1). A slight cross-reactivity between the synthetic proteins was observed in both humoral and cellular tests possibly because of the presence of several copies of the polymerizing sequence Gly-Cys-Cys-Gly in both molecules.

After the seventh day of challenge, the naive receptor and the volunteers who received saline solution had parasitaemias that rose in 12 h from very low levels to >1% (Table 2). Chemotherapy was given immediately with rapid clinical response and without residual effects or sequelae. Two of the four volunteers (H.A. and J.S.) vaccinated with SPf(105)20

showed partial control of the infection with low parasite counts for 13–14 days, followed by development of parasitaemias greater than 0.5%. They were given chemotherapy and recovered with no clinical problems or complications. The other two behaved as the controls and were treated similarly (Table 2). Three of the five volunteers (W.B., W.G. and L.C.) vaccinated with SPf(66)30 had mild infections with steady decrease in parasite counts and total recovery by day 21 (Table 2). Among them, W.B. had a peak of parasitaemia (0.007%) on day 7 but parasites were never seen in his blood again. The other two (W.G. and L.C.) developed parasitaemias below 0.5% that were self-limited by days 18 and 20 post-challenge. Although asexual blood forms had disappeared, a few gametocytes (average 0.004%) were still present. Because very few (less than 0.004%) asexual blood forms were occasionally seen in their blood they were given prophylactic chemotherapy on day 35. By day 40 no gametocytes or asexual blood forms were present and the volunteers' clinical conditions and clinical laboratory results were excellent. The fourth volunteer (D.A.) had a parasitaemia of below 0.45% on day 10, decided to leave the study, and received prompt chemotherapy. The fifth (J.C.) developed parasitaemia similar to the control group (Table 2).

Clinical malaria symptoms (fever, headache, nausea) were present in all volunteers infected with the parasite. Interestingly, the symptoms appeared sooner in the protected individuals than in the controls and naive receptors (day 6). Minor clinical laboratory changes attributable to malaria were observed in all the volunteers but returned to normal levels after chemotherapy. None of the vaccinees developed severe illness and therefore did not require any form of special or intensive care.

Several points of this study deserve emphasis. We used synthetic hybrid molecules containing epitopes from different infectious stages of the parasite. This strategy could be used for future design of vaccines with epitopes from multiple pathogens. The epitopes were polymerized to form large synthetic proteins containing several epitopes, and thus were highly immunogenic, without recourse to foreign carriers^{8,9}. The synthetic proteins were used without special immunopotentiators; addition of such potentiators may further increase their protective capacities. Surprisingly, however, there was no correlation between the different humoral and cellular immunological parameters measured and the anti-malarial protection observed. A considerable delay of the infectious process in half the individuals vaccinated with SPf(105)20 was seen. Because of its design, this molecule may also be effective against sporozoite challenge. SPf(66)30 gave strong protection in most of the inoculated volunteers, even those who received only two doses, suggesting that with more potent adjuvants or slightly improved formulation the protective effect could be close to complete.

Promising results have been obtained by vaccination with the CS protein^{10,11} as well as Pf155/RESA and other molecules⁵. The synthetic protein hybrid polymer SPf(66)30 used here is the first synthetic vaccine for human use against the asexual blood stages of *P. falciparum* malaria.

Supported by the Presidency of Colombia, BID-ICFES Research and Development Program, Ministry of Public Health, the Occidental Petroleum Co. of Colombia and the German Leprosy Relief Association. We thank Drs Belisario-Bentaneur, former President of Colombia, Diego Pizano, Jose Granada, Bruce Merrifield, Richard Houghten and Richard Lerner for encouragement, the Colombian Military Forces especially the volunteers, and the medical staff, and Drs Maria Mercedes Zambrano, Mats Wahlgren, Jose Felix Patino, Mrs Fanny Calvo de Simon, Lida Londoño, Claudia Rocha, Marcelo Rodriguez, Esperanza Rodriguez and Aure Leon for helpful suggestions.

3. Zavala, F. *et al. Science* **228**, 1436–1440 (1985).
4. Good, M. F. *et al. Science* **235**, 1059–1062 (1987).
5. Collins, W. E. *et al. Nature* **323**, 259–262 (1986).
6. WHO *International Digest of Health Legislation* **36**, 825–829 (1985).
7. TDR. Document TDR/IMMAL/FIELDMAL/VAC 85.3.
8. Hersenberg, H. & Tokuhiya, T. *J. exp. Med.* **155**, 1730–1740 (1982).
9. Schutz, M. P., Leclerc, C., Audibert, F. & Chedid, L. *J. Immun.* **135**, 2319–2322 (1985).
10. Ballou, W. R. *et al. Lancet* **i**, 1277–1281 (1987).
11. Herrington, D. A. *et al. Nature* **328**, 257–259 (1987).
12. Barany, G. & Merrifield, R. B., *The Peptides* (eds Gross, E. & Meinhofer, J.) Vol. 2, 1–284 (Academic Press, New York, 1980).
13. Sarin, V. K., Kent, S. B. H., Tam, J. P. & Merrifield, R. B., *Analyt. Biochem.* **117**, 147–157 (1981).
14. Tam, J. P., Heath, W. F. & Merrifield, R. B., *Int. J. Peptide Protein Res.* **21**, 57–65 (1983).
15. Perlmann, H. *et al. J. exp. Med.* **159**, 1686–1704 (1984).
16. Robertson, P. W., Wybin, L. R. & Cox, J. *J. Immun. Meth.* **76**, 195–197 (1985).
17. Troye-Blomberg, M., Perlmann, H. K., Patarroyo, M. E. & Perlmann, P. *Clin. exp. Immun.* **53**, 345–353 (1983).
18. Troye-Blomberg, M., Romero, P., Patarroyo, M. E., Björkman, A. & Perlman, P., *Clin. exp. Immun.* **58**, 380–387 (1984).

Immunological unresponsiveness induced by recipient cells transfected with donor MHC genes

Joren C. Madsen, Riccardo A. Superina*, Kathryn J. Wood† & Peter J. Morris

Nuffield Department of Surgery, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

Immunological unresponsiveness to allografts can be achieved by pretreating recipients with cells (whole blood, erythrocytes, lymphocytes) expressing donor-specific histocompatibility antigens¹. Attempts to determine the relative contribution of class I or class II major histocompatibility antigens and/or minor histocompatibility (miH) antigens towards the induction of an unresponsive state have yielded conflicting results^{2–10}. We have addressed this issue using DNA-mediated gene transfer to introduce murine class I or class II major histocompatibility complex (MHC) genes from the organ donor into cells of recipient origin. This allowed murine recipients of cardiac allografts to be pretreated with syngeneic cells sharing only an isolated class I or class II MHC locus product with the donor organ. We show that pretreatment with either donor class I or class II antigens prolongs survival of cardiac allografts, and that the capacity of a particular donor MHC antigen to induce unresponsiveness is the product of its intrinsic immunogenicity and the antigen load delivered during pretreatment. These results explain discrepancies in earlier studies of antigen-induced unresponsiveness and suggest a novel approach to specific immunosuppression in clinical transplantation.

H-2D, H-2K or H-2IA_α and H-2IA_β genes from C57BL/10 (H-2^b) mice were transfected into C3H/He (H-2^k) derived, class II-negative L-cells and cloned by flow cytometry (Fig. 1). The authenticity of the H-2^b gene product expressed by each clone was confirmed by monoclonal antibody (mAb) binding (Fig. 1), virus-specific killing by H-2K^b or H-2D^b restricted T-cell clones¹¹, and activation of H-2IA^b-restricted T-cells¹². The antigenic site number for each clone was determined by direct antibody saturation binding analysis (Fig. 1). C3H/He (C3H) recipients were intravenously inoculated with either the L-D^b, L-K^b or L-IA^b transfectants one or two weeks before transplantation of a vascularized C57BL/10 (B10) heart graft (Fig. 2).

In control studies, B10 hearts grafted into untreated C3H recipients were consistently rejected in 9.2 days (mean survival time, MST). When C3H recipients were transfused with B10 blood (0.125–0.5 ml) two weeks before transplantation, the MST of B10 hearts was prolonged to 34.2 days (Fig. 2a). This effect was immunologically specific as third-party BALB/c (H-2^d) heart grafts were rejected in 9.7 days (Fig. 2a). When C3H mice

Received 16 September 1987; accepted 8 January 1988.

1. Patarroyo, M. E. *et al. Nature* **328**, 629–632 (1987).

2. Patarroyo, M. E. *et al. in Vaccines 87* (eds Chanock, R. M., Lerner, R. A., Brown, E. & Ginsberg, H.) 117–124 (Cold Spring Harbor Laboratory, New York, 1987).

* Present address: Department of Surgery, Hospital for Sick Children, Toronto, Ontario
† To whom correspondence should be addressed.

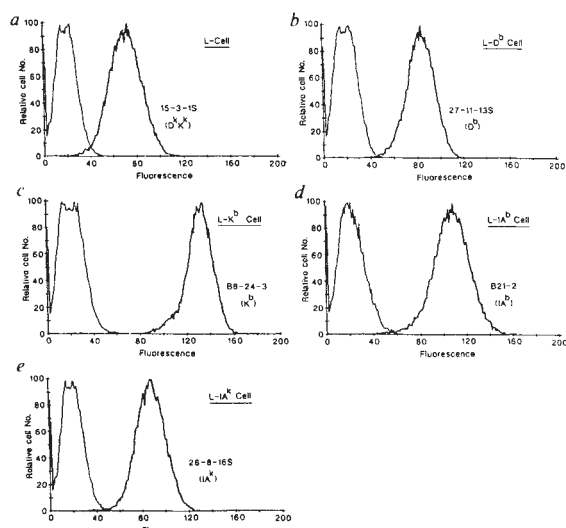


Fig. 1 Surface H-2 antigen expression on parental and transfected L-cell clones. *a*, L-cells *b*, L-D^b cells *c*, L-K^b cells *d*, L-IA^b cells and *e*, L-IA^k cells were stained with a panel of polymorphic anti-H-2 mAbs including B8-24-3 (anti-K^b)²¹, B21-2 (anti-IA^b)²², 27-11-13S (anti-D^b)²³, 26-8-16S (anti-IA^k)²³ and analysed by flow cytometry. All cell lines stained positively with 15-3-1S (anti-D^kK^k)²³ (shown only in *a*). For clarity, all negatively-staining mAbs have been superimposed on the negative control (fluorescein isothiocyanate-labelled affinity purified goat anti-mouse antibody alone). Clone L-D^b R-1, expressed $1.0 \pm 0.1 \times 10^5$ (mean $\pm$ sd) D^b molecules per cell, clone L-K^b G, $2.2 \pm 0.3 \times 10^5$ K^b molecules per cell and clone L-IA^b 1-C, $7.7 \pm 1.8 \times 10^5$ IA^b molecules per cell. **Methods.** The L-D^b clone R-1 was generated by transfecting neomycin-sensitive L-cells with the cosmid clone B 1.28, containing the H-2D^b gene and neomycin-resistance gene and selecting with G418 (Geneticin 0.5 mg ml⁻¹, Gibco). The L-K^b clone, G-1 was subcloned from LH8²⁴. The L-IA^b clone, 1-C and L-IA^k clone 1-A were subcloned from NABBIF¹² and RT7.3H3M²⁵ respectively. All lines were mycoplasma-negative and were grown in selection media containing either HAT (hypoxanthine 15 μ g ml⁻¹, aminopterin 0.2 μ g ml⁻¹ and thymidine 5 μ g ml⁻¹), G418, or HAT plus MX (mycophenolic acid 6 μ g ml⁻¹ and xanthine 250 μ g ml⁻¹). For analysis by flow cytometry, adherent transfectants were trypsinized with 0.125% trypsin (w/v) and left in suspension overnight. Cells (5×10^5) were incubated with the indicated mAb (as culture supernatant) in conditions of antibody excess for 1 hr at 4°C. Cells were washed, then incubated for an additional hour at 4°C with either fluoresceinated, affinity purified, goat F(ab')₂ anti-mouse IgG (Sigma) or fluoresceinated, affinity purified, goat anti-rat IgG (Sigma). After two more washes the cells were analysed using a Orthocytofluorograf System 50 (Ortho Diagnostic Systems) with log amplification. To determine antigenic site numbers for the L-D^b, L-K^b and L-IA^b transfectants, purified, high affinity mAbs 27-11-13S, B8-24-3 and B21-2 were labelled with ¹²⁵I by the chloramine T method. Recovery of mAb was 60–70% with >65% of the labelled antibody retaining antigen specificity. The ¹²⁵I-labelled antibody was then incubated with the appropriate transfectant at 4°C for 90 min. Binding assays met the requirements of ligand specificity and specific binding saturability. After washing away unbound antibody, the number of antibody molecules bound per cell was calculated from the specific c.p.m. bound and the known specific activity of the radiolabelled reagent. This number was used to determine the number of antigenic sites per cell. Assay conditions were chosen so that monoavalent binding of mAb would predominate²⁶.

were treated with syngeneic, parental L-cells at various doses either one or two weeks before transplantation, allograft survival was not prolonged (Fig. 2a). Furthermore, pretreatment with L-cells transfected with an irrelevant chimaeric influenza haemagglutinin gene (clone DWK-17) or recipient-specific IA^k _{α} and IA^k _{β} genes (clone L-IA^k 1-C) did not prolong allograft survival (Fig. 2b). Thus, neither the antigens constitutively expressed by parental L-cells, nor L-cells expressing recipient-

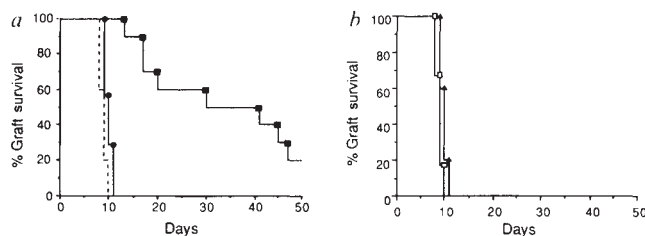


Fig. 2 Survival curves of cardiac allografts transplanted into C3H/He mice pretreated with either donor-specific whole blood or different L-cell preparations. Two weeks before transplantation C3H recipients were given a single intravenous injection of 0.125–0.5 ml of B10 whole blood, then grafted with either a B10 heart (—■—) (MST = 34.2 days, range = 13–72 days; *n* = 14) or a third-party BALB/c heart (----) (MST = 9.7 days, range = 8–10 days; *n* = 5). C3H recipients were also pretreated two weeks before grafting with either *a*, 10^7 untransfected L-cells (—●—) (MST = 9.8 days, range = 9–11 days; *n* = 7) or *b*, 10^7 DWK-17 cells (L-cells transfected with a chimaeric influenza haemagglutinin gene; —□—) (MST = 8.8 days, range = 8–10 days; *n* = 6) or 10^6 L-IA^k transfectants (—▲—) (MST = 9.8 days, range = 9–11 days, *n* = 5). Untreated C3H recipients rejected both a B10 graft and BALB/c graft between eight and ten days after transplantation (B10 heart MST = 9.2 ± 0.75 days, BALB/c heart MST = 8.75 ± 0.6 days). Not shown are B10 graft survival times after C3H recipients received identical pretreatment inoculums one week before grafting or were pretreated either one or two weeks before transplantation with 2.5×10^6 or 5×10^6 L-cells, 2.5×10^6 or 5×10^6 DWK-17 cells and 5×10^6 or 10^7 L-IA^k cells, where all grafts were rejected in a similar manner to untreated recipients.

Methods. Confluent cell monolayers were trypsinized, then allowed to recover in suspension overnight. Cells were resuspended in heparinized (15 units per ml) 0.9% saline and injected into the recipient's dorsal penile vein. Cell viability was 95–100%. Seven or 14 days later, heart grafts from male B10 donors were heterotopically transplanted into the abdomen of the pretreated C3H recipients using microsurgical techniques²⁷. Allograft rejection was defined as the complete loss of palpable cardiac contractions and/or the cessation of all electrical activity on electrocardiogram²⁸. Rejection was confirmed histologically. Survival times between experimental and control groups were compared using the Wilcoxon rank sum test.

specific IA^k surface molecules, nor non-specific effects resulting from the transfection process itself, altered the host's ability to reject a vascularized B10 cardiac allograft.

To evaluate the role of donor-specific H-2D locus products, C3H recipients were treated intravenously with 10^7 L-D^b cells (clone R-1) two weeks before heart grafting. This resulted in a modest but significant donor-specific prolongation of allograft survival (MST = 14.3 days, Fig. 3a). But larger (2×10^7 cells) or smaller ($\leq 5 \times 10^6$ cells) doses of L-D^b cells were ineffective (Fig. 3b).

Next, C3H recipients were treated with 10^7 L-K^b cells (clone G) two weeks before transplantation. This specifically prolonged B10 allograft survival to 14.4 days (Fig. 3d), similar to that observed with 10^7 L-D^b cells. L-K^b clone G, however, expressed approximately twice the number of surface H-2^b molecules as the L-D^b clone R-1 (Fig. 1), so to deliver an equivalent load of H-2K^b antigen, the pretreatment dose was halved. Reducing the L-K^b inoculum to 5×10^6 cells dramatically improved the mean survival time to 45.2 days, with three of twelve allografts going on to indefinite survival (>100 days, Fig. 3c). The dose response to the L-K^b cells is shown in Fig. 3d.

Thus, in contrast to earlier studies^{2,6,8}, pretreatment with donor-specific MHC class I locus products specifically prolongs allograft survival in the absence of donor class II or mI antigens. But this effect is critically dependent upon the total antigen load delivered to the recipient. Indeed, increasing the dose of donor-specific B10 blood used to pretreat C3H recipients in Fig. 2a (0.25 ml–0.5 ml) to 1 ml, also resulted in a decrease in MST from 34.2 days to 15.4 days. Failure to induce allograft

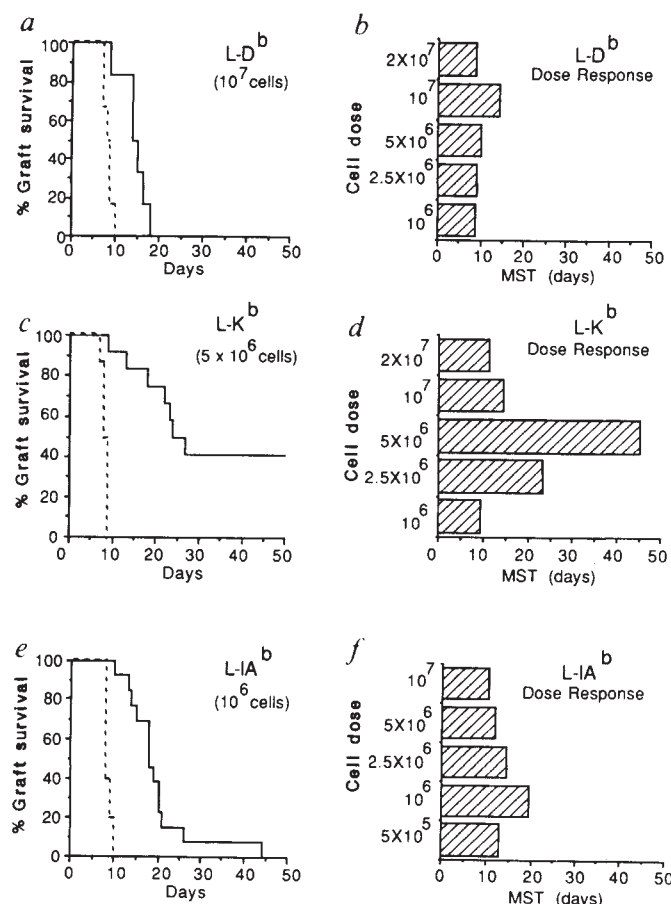


Fig. 3 Survival of cardiac allografts in C3H recipients pretreated with L-cell clones transfected with donor-specific class I or class II genes. *a*, *c* and *e* show graft survival against time for either B10 hearts (—) or third party BALB/c hearts (---) grafted into C3H recipients, pretreated with optimal doses of *a*, L-D^b *c*, L-K^b or *e*, L-IA^b transfectants. Cell doses are in parentheses. The corresponding plots of cell dose versus mean survival time for the B10 to C3H combination are shown in *b* (L-D^b cells), *d* (L-K^b cells) and *f* (L-IA^b cells). After L-D^b pretreatment, prolongation of B10 cardiac allograft survival reached statistical significance with 10⁷ cells; $P < 0.05$, MST = 14.3 days, range = 9–18 days, ($n = 6$). After L-K^b pretreatment, significant prolongation of graft survival was obtained with 10⁷ cells; $P < 0.05$, MST = 14.4 days, range = 10–18 days, ($n = 7$), 5 × 10⁶ cells; $P < 0.01$, MST = 45.2 days, range = 9–>100 days, ($n = 12$) and 2.5 × 10⁶ cells; $P < 0.01$, MST = 23.2 days, range = 9–35 days, ($n = 5$). L-IA^b cell pretreatment resulted in significant allograft prolongation with 2.5 × 10⁶ cells; $P < 0.05$, MST = 14.7 days, range = 10–22 days, ($n = 9$), and 10⁶ cells; $P < 0.01$, MST = 19.6 days, range = 10–44 days, ($n = 13$).

unresponsiveness by pretreatment with supraoptimal doses of donor class I antigen has been observed in other allogeneic systems^{9,10} where it has been attributed to recipient sensitization¹⁰.

At equivalent class I antigen loads, pretreatment with isolated donor-specific H-2K locus antigen prolonged allograft survival more effectively than isolated H-2D products. Furthermore, C3H recipients, subcutaneously injected with equal antigenic doses of transfected L-cells, rejected B10 hearts more rapidly following L-K^b sensitization than L-D^b sensitization (data not shown). Similar immunodominance of the K locus has been observed in murine models of skin graft rejection¹³, neonatal tolerance¹⁴ and tumour enhancement¹⁵. Taken together, these data suggest that the induction of unresponsiveness with a specific class I antigen is related to its strain-specific immunogenicity.

Pretreatment with cells expressing donor-specific class II anti-

gen has been shown to prolong allograft survival in some studies⁴ and sensitize in others⁵. We predicted that the optimal effect of L-IA^b pretreatment (sensitizing or enhancing) would occur at a lower cell dose than with the class I transfectants, because the L-IA^b clone 1-A expressed 3–4 times as much H-2^b surface antigen. Fig. 3*e, f* confirms that pretreatment with 10⁶ or 2.5 × 10⁶ L-IA^b cells administered one week before grafting resulted in donor-specific prolongation of B10 allograft survival. Supraoptimal doses (5 × 10⁶ cells and 10⁷ cells) neither delayed nor accelerated the rejection response. These data suggest that the capacity of IA-bearing cells to act as sensitizing alloimmunogens is not conferred merely by the expression of class II MHC antigen¹⁶.

Thus, using a method of alloantigen presentation which eliminates (1) Ir gene effects, (2) contamination by extraneous H-2 and/or miH antigens, (3) variation in antigen load, (4) immunologically specialized lymphohaemopoietic cells and (5) the need for concurrent immunosuppression, we have reconciled discrepancies in the literature^{2–10} by showing that class I and class II MHC subregion products can each induce specific unresponsiveness; but that the tolerizing potential of a particular donor-specific antigen is a function of its intrinsic immunogenicity (strain-specific immunodominance) and the antigenic content of the pretreatment inoculum (antigen load).

How does pretreatment with a single donor H-2 gene product induces unresponsiveness to an organ allograft which is incompatible for the entire H-2 complex as well as multiple miH antigens? Possibly, the immune response to an allograft represents a hierarchy of responses to incompatible MHC and miH antigens, and suppressing the response to one immunodominant antigen allows the response to other less immunogenic antigens to be overridden. This would explain the observed hierarchy of MHC antigens in the induction of unresponsiveness (K locus > IA locus ≥ D locus), and the results of our preliminary studies, showing that pretreatment of recipients with combinations of donor-specific H-2 molecules gives additive effects, extending the dose range that shows graft prolongation. The mechanism of induction of unresponsiveness is still unknown, although specific suppressor cells primed by antigen pretreatment might suppress responses to other alloantigens carried on the same cell¹⁷. Indeed, in adoptive transfer experiments, suppressor cells were detected in the spleens of L-K^b pretreated recipients bearing long-term surviving allografts, whereas serum had no effect suggesting that alloantibody did not mediate this effect (data not shown).

Our results support the hypothesis that the effectiveness of random blood transfusions in improving graft survival in clinical renal transplantation¹⁸ is due to fortuitous sharing of one or more MHC antigens between the blood donors and the cadaveric organ donor¹⁹.

Finally, the fact that a single class I locus antigen can be more effective than whole blood in enhancing allograft survival (compare Fig. 2*a* and 3*c*) suggests that the full potential of antigen pretreatment in clinical transplantation has not been realised. Possibly, monoclonal populations of transfected mammalian cells expressing enhancing HLA antigens could be used to induce specific immunosuppression to organ grafts (purified, soluble antigen alone is ineffective²⁰), diminishing the chance of recipient sensitization and eliminating the risk of hepatitis and AIDS.

We thank Dr Ronald Germain for cell lines NABBIF and RT7.3H3M, Dr Andrew Mellor for cosmid B1.28 and cell line LH8 and Professor D. W. Kingsbury for cell line DWK-17. We thank Drs Margaret Dallman, Elizabeth Simpson and Alain Townsend for critical readings of the manuscript, Dr Donald Mason and Dr Robert Lechler for helpful discussion and Mr Gareth Plant for technical assistance. This work was supported by the Medical Research Council, UK, and the British Heart Foundation. J.C.M. is a Claude E. Welch Surgical Research Fellow, Massachusetts General Hospital, Boston.

Received 13 November 1987; accepted 2 February 1988.

- Medawar, P. B. *Transplantation* 1, 21-38 (1963).
- Majoor, G. D. & van Breda Vriesman, P. J. C. *Transplantation* 41, 42-96 (1986).
- Wood, K. J. & Morris, P. J. *Transplantation* (in the press).
- Nagata, M. *et al.* *Transplantation* 38, 522-526 (1984).
- Faustman, D. L., Lacy, P. E., Davie, J. M. & Hauptfeld, V. *Transplant. Proc.* 15, 1341-1343 (1983).
- Wood, M. L. & Monaco, A. P. *Transplantation* 37, 35-42 (1984).
- Guttmann, R. D., Forbes, R. D. C., Cramer, D. V. & Gill, T. J. *Transplantation* 30, 216-218 (1980).
- Soulillou, J. P., Blandin, F., Günther, E. & Lemoine, V. *Transplantation* 38, 63-67 (1984).
- Wood, K. J., Evins, J. & Morris, P. J. *Transplantation* 39, 56-61 (1985).
- Wood, P., Horsburgh, T. & Brent, L. *Transplantation* 31, 8-14 (1981).
- Townsend, A. R. M., Taylor, P. M., Mellor, A. L. & Askonas, B. A. *Immunogenetics* 17, 283-294 (1983).
- Lechler, R. I., Ronchese, F., Braunstein, N. S. & Germain, R. N. *J. exp. Med.* 163, 678-696 (1986).
- McKenzie, I. F. C. & Henning, M. M. *J. Immunogen.* 4, 249-257 (1977).
- Streilein, J. W. & Klein, J. *J. Immun.* 119, 2147-2150 (1977).
- Wallich, R. *et al.* *Nature* 315, 301-305 (1985).
- Lechler, R. I. & Batchelor, J. R. *J. exp. Med.* 155, 31-41 (1982).
- Hutchinson, I. V. *Transplantation* 41, 547-555 (1986).
- Opelz, G. *et al.* *Transplantation Proc.* 17, 2357-2361 (1985).
- Hutchinson, I. V. & Morris, P. J. *Transplantation* 41, 166-169 (1986).
- Spencer, S. C. & Fabre, J. W. *Transplantation* 44, 141-148 (1987).
- Köhler, G. *et al.* *Immun. Syst.* 2, 202-208 (1980).
- Steinman, R. M. *et al.* *J. exp. Med.* 159, 1248-1255 (1980).
- Ozato, K. & Sachs, D. H. *J. Immun.* 126, 317-321 (1981).
- Mellor, A. L. *et al.* *Nature* 298, 529-534 (1982).
- Lechler, R. I., Norcross, M. A. & Germain, R. N. *J. Immun.* 135, 2914-2922 (1985).
- Mason, D. W. & Williams, A. F. *Biochem. J.* 187, 1-20 (1980).
- Corry, R. J., Winn, H. J. & Russell, P. R. *Transplant. Proc.* 5, 733-735 (1973).
- Superina, R. A., Peugh, W. N., Wood, K. J. & Morris, P. J. *Transplantation* 42, 226-227 (1986).

Haemophilia A resulting from *de novo* insertion of *L1* sequences represents a novel mechanism for mutation in man

Haig H. Kazazian Jr, Corinne Wong,
Hagop Youssoufian*, Alan F. Scott†,
Deborah G. Phillips & Stylianos E. Antonarakis

Genetics Unit, Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA
† Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

L1 sequences are a human-specific family of long, interspersed, repetitive elements, present as $\sim 10^5$ copies dispersed throughout the genome¹. The full-length *L1* sequence is 6.1 kilobases, but the majority of *L1* elements are truncated at the 5' end, resulting in a fivefold higher copy number of 3' sequences¹. The nucleotide sequence of *L1* elements includes an A-rich 3' end and two long open reading frames (*orf-1* and *orf-2*), the second of which encodes a potential polypeptide having sequence homology with the reverse transcriptases¹⁻⁴. This structure suggests that *L1* elements represent a class of non-viral retrotransposons^{1,2}. A number of *L1* complementary DNAs, including a nearly full-length element, have been isolated from an undifferentiated teratocarcinoma cell line⁵. We now report insertions of *L1* elements into exon 14 of the factor VIII gene in two of 240 unrelated patients with haemophilia A. Both of these insertions (3.8 and 2.3 kilobases respectively) contain 3' portions of the *L1* sequence, including the poly (A) tract, and create target site duplications of at least 12 and 13 nucleotides of the factor VIII gene. In addition, their 3'-trailer sequences following *orf-2* are nearly identical to the consensus sequence of *L1* cDNAs (ref. 6). These results indicate that certain *L1* sequences in man can be dispersed, presumably by an RNA intermediate, and cause disease by insertional mutation.

Haemophilia A is an X-linked disorder of blood coagulation caused by deficiency of factor VIII. The factor VIII gene has been cloned and characterized, and a number of gene defects

producing haemophilia A have been identified⁷⁻¹¹. We have screened the DNA of 240 unrelated males with haemophilia A for any abnormalities in the factor VIII gene detectable by restriction analysis. Significant deletions were found in 15 patients and point mutations in 12 (ref. 11 and unpublished results). In addition, the DNA of two unrelated patients had insertions in exon 14 of the factor VIII gene. In patient 1 (JH-27), exon 14 fragments were increased in size by 3.8 kilobases (kb) and in patient 2 (JH-28) by 2.3 kb (Fig. 1). No abnormalities were detected in the factor VIII genes in the parents of both patients (Fig. 1), suggesting that in both cases the mutations had arisen *de novo*, although maternal mosaicism cannot be excluded. In both patients the disease was severe and no factor VIII activity was detectable.

To characterize these insertions, abnormal 3.5-kb and 6.5-kb *EcoRI* fragments containing the 5' and 3' ends of the insertion of JH-27, and the abnormal 6.9-kb *EcoRI* fragment containing the entire insertion of JH-28, were cloned in λ gt10 (ref. 12). Pertinent fragments containing the breakpoints and internal portions of the insertions were subcloned and sequenced¹³ (Fig. 2).

The insertions of JH-27 contained a 3.8-kb truncated *L1* sequence from nucleotide 2,363 of the genomic *L1* consensus sequence² to the 3' end at nucleotide 6,161, followed by a 57-nucleotide tract of A residues. The insertional event resulted in duplication of at least 12 nucleotides of the target site in the factor VIII gene from nucleotides 3,054 to 3,065 of the factor VIII cDNA sequence. Interestingly, this 12-nucleotide target site is rich in A residues on the coding strand (8 of 12 residues) (Fig. 2).

The insertion in JH-28 also included the 3' end of an *L1* sequence (nucleotides 4,020 to 6,161), but it is more complex (Fig. 2). This insertion contained two blocks of sequence, from nucleotides, 4,020-5,114 and 5,115 to 6,161 in a head-to-head arrangement. Like the insertion in JH-27, the block of sequence 5,115-6,161 contained a highly A-rich 3' sequence of 77 nucleotides, the last 30 of which were A residues. Although this *L1* insertion is rearranged, a clean break between nucleotides 5,114 and 5,115, a polarity change and gap repair occurred without loss of *L1* or factor VIII sequences. The *L1* insertion also created a target site duplication of at least 13 nucleotides of the factor VIII sequence from nucleotides 3,653-3,665 of the factor VIII cDNA sequence. Again, the target site sequence is A-rich in the coding strand (9 of 13 nucleotides).

A consensus sequence for the 3'-trailer region of *L1* cDNA has been published⁶. This sequence, obtained after analysis of 19 *L1* cDNA isolated separately, differs in 15 positions out of 186 nucleotides from the consensus sequence derived from human genomic *L1* elements². We have compared the 186 3' nucleotides of our two inserted *L1* sequences to both the cDNA and genomic consensus sequences. The inserted sequences differ from each other by only a single nucleotide in this region, and they have the cDNA consensus nucleotide changes at 14 of 15 sites (Fig. 3). Furthermore, 4 of the 19 cDNAs (subset Ta) differ from the cDNA consensus at four additional nucleotides (Fig. 3). The insert of JH-27 had these four additional changes, and the insert of JH-28 had three of the four changes. Both insertions contained a G residue, 11 nucleotides upstream of the A-rich region. This G has been observed in one of the published cDNA sequences which has the same sequence in the 3'-trailer region as the insert of JH-27. These data provide evidence that our *L1* insertions are the result of invasion of an RNA intermediate or of cDNA into the coding sequences.

These *L1* insertions are the first large non-viral insertions described in man which are not due to expansion of short repeats by unequal crossing-over events. A small segment of an *L1* sequence (36-41 nucleotides) has been found between the breakpoints of a deletion involving the β -globin gene cluster¹⁴. Insertions of species-specific elements related to human *L1* elements have been observed in other organisms. An analogous

* Present address: Hematology-Oncology Unit, Cox 6, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA.

Fig. 1 Restriction endonuclease analysis of DNA from pertinent members of families JH-27 and JH-28. The probe is a 3' factor VIII cDNA fragment spanning exons 14-26. In the left panel, *TaqI* digest of DNA from JH-27 shows loss of the major 5.9-kb fragment, which includes exon 14, and its replacement with fragments of 5.7 kb and 4.0 kb. The minor fragment (~5.9 kb) which remains in patients DNA is a normal fragment derived from exons 20 to 22. Further analysis showed insertion of nucleotides 2,363-6,161 of an *L1* element, plus an A-rich tract, into exon 14 (Fig. 2). Two new *TaqI* fragments are expected from patient DNA as the *L1* consensus sequence in this region contains a single *TaqI* site at nucleotide 3,415. Note that DNA of mother and maternal grandmother lack the abnormal 5.7-kb and 4.0-kb fragments. In the right-hand panels are *SstI* and *KpnI* digests of JH-28 and his mother, probed with the 3' factor VIII cDNA. In JH-28 the major 3.2-kb *SstI* fragment of exon 14 is replaced by an abnormal 5.5-kb fragment. The 5.5-kb fragment is not seen in maternal DNA. Also in the patient DNA, the normal 7.3-kb *KpnI* fragment containing exon 14 is replaced by 5.3-kb and 4.3-kb fragments. These abnormal fragments are not present in maternal DNA. Further analysis showed that JH-28 had insertion into exon 14 of nucleotides 4,020-6,161 of an *L1* element, plus an A-rich tract. This region of the *L1* consensus sequence lacks an *SstI* site, and the nucleotide sequence of the *L1* element from JH-28 contains a single *KpnI* site at nucleotide 4,786, thereby explaining the fragment patterns observed.

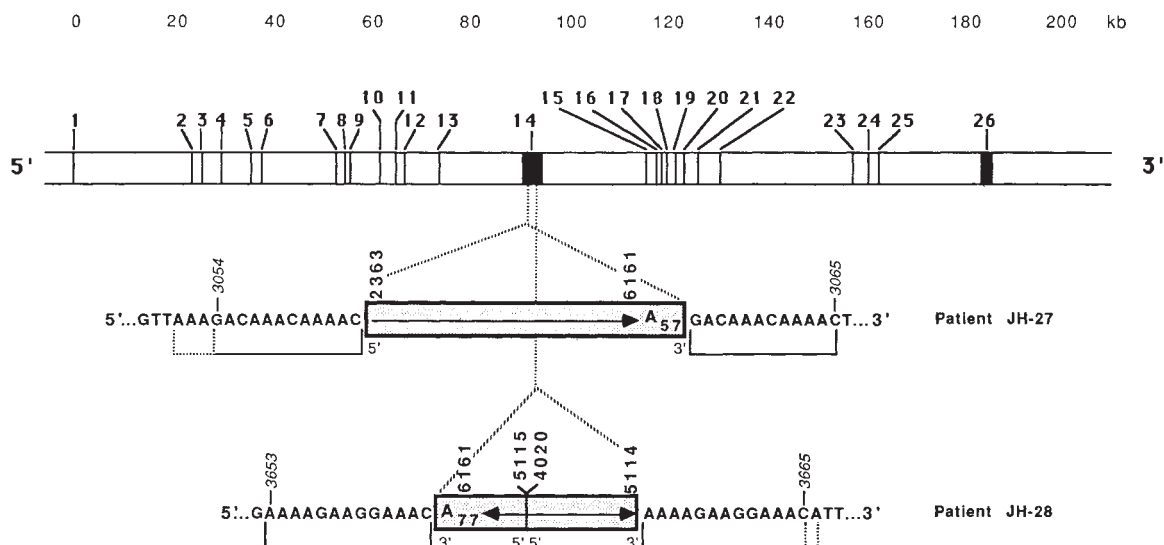
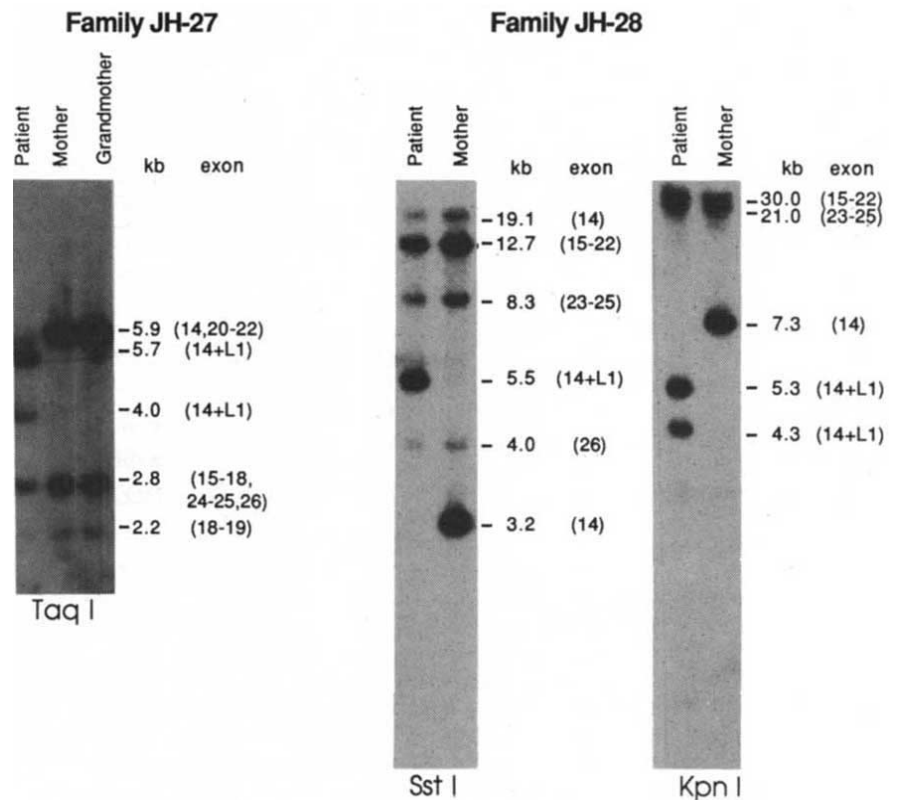


Fig. 2 Diagram of *L1* insertions in exon 14 of the factor VIII gene. The factor VIII gene spans 186 kb and contains 26 exons. In patients JH-27 and JH-28, restriction analysis of leukocyte DNA with a cDNA probe spanning exons 14-26 gave hybridizing fragments of increased size. After cloning the abnormal *EcoRI* fragments derived from exon 14 in λ gt10 (ref. 12) and demonstrating that these *EcoRI* fragments contained 3' *L1* sequences by hybridization to an *L1* probe²⁰, the cloned fragments were subcloned into M13 and partially sequenced¹³. The 3.8-kb *L1* insertion from patient JH-27 is flanked by a 12-base pair target site duplication of factor VIII cDNA sequence (nucleotides 3,054-3,065, where nucleotide 1 is A of the initiator codon)⁸. Residues 3,051-3,053 of the factor VIII cDNA are adenylic acids and could also be duplicated (shown by the hatched bracket). The rearranged 2.3-kb insertion from patient JH-2 is shown and is flanked by a 13-base pair target site duplication. Residue 3,666 of the factor VIII cDNA is an A and could also be duplicated. Filled boxes represent the *L1* elements, and the arrows within the boxes point towards the 3' end of the *L1* sequence. The *L1* insertion of patient JH-27 was sequenced between nucleotides 2,363-2,551, 4,964-5,178, 5,977-6,161 and the poly(A) tract, and its restriction map (using *EcoRI*, *BamHI*, *SstI* and *KpnI*) was identical to the restriction map of the *L1* genomic consensus sequence². The *L1* insertion of patient JH-28 was sequenced between nucleotides 4,020-4,237, 4,542-5,269, 5,613-6,161 and the poly(A) tract. Both *L1* insertions showed 98% similarity to the consensus genomic *L1* sequence outside of the 3'-trailer region.

L1 Genomic	...	ACGAACGGGAACATCACACACTGGGGCCTGTTCGGGTGGGGGNGGGGGAGGCATACGA	6032
L1 cDNA		T T A G C A	
JH-27 Insert		T T A G C A	
JH-28 Insert		T T A G C A	

Fig. 3 Comparison of 3' trailer region sequence (186 nucleotides) of the L1 element from genomic consensus² (top line), cDNA consensus⁶ (middle line) and factor VIII L1 insertion sequences of JH-27 and JH-28. The cDNA consensus shown is that of subset Ta (ref. 6), and the additional four nucleotides which differ from the L1 cDNA and genomic consensus sequences are underlined.

TTAGGAGATATACCTAATGCTAAATGACAGTAAATGGTGCAGCACCAACATGGGACAT	6092
G G A G G G G	
G G A G G G G	
G G A G G G G	
GTATACATATGTAACAAACCTGCACGTTGTCACATGTACCTAGAACTTAACTATAATAA	6152
T AA A A	
T AA A G	
T AA A G	

spontaneous insertion of an F element in *Drosophila*, which has a similar organization and sequence to the L1 elements of mammals, produces a reversion of the white-ivory mutation at the white locus¹⁵. F-element insertion sites contain 8–13 nucleotide duplications which are not A-rich, but are random in sequence¹⁶. The presence of a canine L1 homologue 5' to a *myc* gene has been implicated in the aetiology of canine venereal tumour¹⁷.

The consensus sequences of mammalian L1 elements and *Drosophila* F elements contain two open-reading frames (*orf-1* and *orf-2*) (refs 2–4, 18, 19). *Orf-2* has blocks of sequence similarity to the polymerase domain of various reverse transcriptases^{3,4}. Because full-length L1 RNAs have been detected in the poly(A)⁺ cytoplasmic RNA of human teratocarcinoma cells, it has been postulated that in germ-line cells a small number of full-length L1 sequences are transcribed and translated¹. The polypeptide of *orf-2* has been implicated as a reverse transcriptase of the L1 RNA itself in producing cDNA copies, many of which are truncated. These cDNAs could then re-enter genomic DNA, perhaps at A-rich sequences, as implied by the sequences of the target sites in the factor VIII gene. We postulate that the poly(T) tail of L1 cDNA is involved in base-pairing with the A-rich factor VIII sequences after staggered single-strand breaks in the factor VIII gene. Rejoining of the L1 cDNA to the factor VIII sequence and filling-in of the complementary strand of the cDNA would complete the event.

Insertion of L1 elements, involving retrotransposition of DNA sequences through an RNA intermediate into a new and distant location in the genome, represents a mechanism for mutation to produce human disease fundamentally different from those previously described. Because we do not know when these L1 insertion events occur, whether in the sperm or ovum, after fertilization, or during early stages of embryogenesis, the proportion of such insertions that are heritable is unknown. Yet finding two L1 insertions among 240 patients with haemophilia A suggests that this mechanism of mutation is not uncommon.

We thank Drs T. Fanning, M. Singer and M. Edgell for helpful discussions, E. Pasterfield for assistance in preparation of the manuscript and J. Strayer for the artwork. H.H.K., A.F.S. and S.E.A. are supported by grants from the NIH.

Received 11 November 1987; accepted 22 January 1988.

- Fanning, T. & Singer, M. F. *Biochem. biophys. Acta* (in the press).
- Scott, A. F. *et al. Genomics* **1**, 113–125 (1987).
- Hattori, M. *et al. Nature* **321**, 625–628 (1986).
- Sakaki, Y. *et al. Cold Spring Harb. Symp. quant. Biol.* **51**, 465–469 (1986).
- Skowronski, J. & Singer, M. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6050–6054 (1985).
- Skowronski, J. & Singer, M. F. *Cold Spring Harb. Symp. quant. Biol.* **51**, 457–464 (1986).
- Gitschier, J. *et al. Nature* **312**, 326–330 (1984).
- Toole, J. J. *et al. Nature* **312**, 342–347 (1984).
- Gitschier, J. *et al. Nature* **315**, 427–430 (1985).
- Yousoufian, H. *et al. Nature* **324**, 380–382 (1986).
- Antonarakis, S. E., Yousoufian, H. & Kazanian, H. H. Jr *Molec. biol. Med.* **4**, 81–94 (1987).
- Maniatis, T., Fritsch, E. G. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor, 1982).
- Sanger, K., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Mager, D. L., Henthorn, P. S. & Smithies, O. *Nucleic Acids Res.* **13**, 6559–6575 (1985).
- DiNocera, P. P. & Casari, G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5843–5847 (1987).
- DiNocera, P. P., Digan, M. E. & Dawid, I. B. *J. molec. Biol.* **168**, 715–727 (1983).
- Katzir, N. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 1054–1058 (1985).
- Loeb, D. D. *et al. Molec. cell. Biol.* **6**, 168–182 (1986).
- Soares, M. B., Schon, B. & Elstratidis, A. *J. molec. Evol.* **22**, 117–133 (1985).
- Adams, J. W. *et al. Nucleic Acids Res.* **8**, 6113–6128 (1980).

Oncogene *jun* encodes a sequence-specific *trans*-activator similar to AP-1

Peter Angel, Elizabeth A. Allegretto, Steve T. Okino, Kazue Hattori, William J. Boyle*, Tony Hunter* & Michael Karin

Department of Pharmacology M-036, Center for Molecular Genetics, School of Medicine, University of California, San Diego, La Jolla, California 92093, USA

* The Salk Institute for Biological Sciences, San Diego, California 02138, USA

Proto-oncogenes encode proteins with three main sites of action: the cell-surface membrane, the cytoplasm and the nucleus^{1,2}. Although the exact biochemical function of most proto-oncogene products is not understood, several of them are known to be involved in signal transduction^{3–7}. A role in gene regulation through DNA binding has been suggested for a recently isolated member of the group of oncogenes acting at the nucleus, *v-jun*. The C-terminus of the putative *v-jun*-encoded protein is similar in sequence to the C-terminus of the yeast transcriptional activator GCN4 (refs 8, 9), which forms its minimal DNA-binding domain¹⁰. GCN4 binds to specific sites whose consensus sequence¹¹ is highly similar to the recognition sequence of the mammalian transcriptional activator AP-1 (refs 12, 13). Like GCN4, AP-1 binds to promoter elements of specific genes and activates their transcription^{12–15}. Because of the similarity between the recognition sites for GCN4 and AP-1, we examined the possibility that AP-1 could be the product of the *c-jun* proto-oncogene. The experimental results reported here indicate that the JUN oncoprotein is a sequence-specific transcriptional activator similar to AP-1.

Recently AP-1 has been purified from HeLa cells as a protein of relative molecular mass (*M_r*) 44,000–45,000 (44–45K)^{12,13}. Mapping of tryptic peptides indicates that two bands of *M_r* 44–45K and 40K present in highly purified preparations of AP-1 (see Fig. 1, lane 4) are structurally related (F. Mercurio and W.J.B., unpublished results). To examine the relationship between AP-1 and *v-JUN* we tested whether antisera raised against two specific peptides derived from the *v-JUN* sequence¹⁶ can interact with AP-1. Different AP-1 preparations were transferred to nitrocellulose membranes and incubated with the anti-peptide antisera (kindly provided by T. Bos and P. Vogt). As shown in Fig. 1, antiserum directed against a synthetic peptide (USC 4) derived from the putative DNA-binding domain of *v-JUN* (amino acids 199–215; see Fig. 2) reacted with polypeptides having the same mobility as AP-1 (see lanes 2–5). AP-1 reacted in the same way with another antiserum raised against an unrelated peptide (USC 2) whose sequence is located within the N-terminal portion of *v-JUN* (amino acids 75–87; see lane

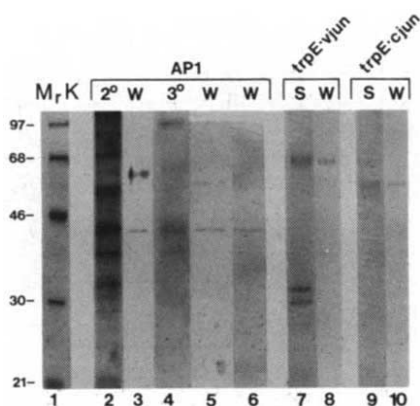


Fig. 1 Immunoblot analysis of AP-1 and trpE-JUN fusion proteins. Lanes 2 and 3, AP-1 (25 ng) after a second pass on the affinity column; lanes 4–6, AP-1 (5 ng) after a third pass on the affinity column; lanes 7 and 8, Bp 55^{trpE-cjun} (5 ng); lanes 9 and 10, Bp 75^{trpE-vjun} (5 ng), subjected to SDS-PAGE and Western blotting. Lanes 1, 2, 4, 7 and 9 show the silver-stained proteins (labelled M_r 2°, 3° or S) and lanes 3, 5, 8 and 10 (labelled W) show immunoblots with anti-USC4 antiserum, and lane 6, the reaction with anti-USC2 antiserum. Protein was estimated by comparison with known standards (lane 1). The strong upper immunoreactive bands in lane 3 are an artefact due to reaction with dithiothreitol in the sample buffer and do not correspond to protein (compare lane 2), being seen with sample buffer alone (not shown), so dithiothreitol was subsequently omitted.

Methods. HeLa whole cell extracts were chromatographed on heparin agarose, Sephacryl S-300 and then on a sequence specific DNA-affinity resin for AP-1, as described¹². TrpE-JUN fusion protein extraction and preparation are described in legend to Fig. 4. AP-1 (affinity-purified and precipitated with trichloroacetic acid) and fusion-protein extracts were subjected to SDS-PAGE (ref. 12) and immunoblot (Western) analysis²⁷ using an alkaline phosphatase-conjugated anti-rabbit IgG colour development kit (BioRad).

6). No significant cross reaction was observed between these antisera and purified AP-2 or the crude S-300 fraction, which contains a very low concentration of AP-1. In addition, no reaction was observed using pre-immune sera (data not shown). These results indicate that AP-1 shares at least two different epitopes with v-JUN.

To investigate these structural similarities further, a human *c-jun* complementary DNA clone was isolated from a human fibroblast cDNA library using two synthetic oligodeoxynucleotide probes derived from the evolutionally conserved putative DNA-binding domain of the v-jun-encoded proteins. Sequencing of this clone, hcJ-1, showed a long open-reading frame (ORF), predicting an amino acid sequence highly similar to that of v-JUN (Fig. 2). The ORF is followed by a large 3'-untranslated region containing two polyadenylation signals.

Hybridization of probes derived from hcJ-1 to poly(A)⁺ RNA from HeLa cells revealed transcripts of ~2.6–2.7 and 3.4 kilobases (kb) (data not shown). The same transcripts were observed using probes from the 5' and 3' regions of hcJ-1, suggesting that they are derived from the same gene. We propose that the 2.6–2.7-kb transcripts are the mature *c-jun* transcripts, which differ in size because of two alternative polyadenylation sites in hcJ-1, and that the 3.4-kb transcript is probably the primary unspliced *c-jun* transcript. A single fragment in digests of human genomic DNA strongly hybridizes to either of these probes, which also suggests that all the transcripts are derived from a single gene. There are several weak cross-hybridizing bands, so possibly other genes less homologous to *c-jun* are present (data not shown).

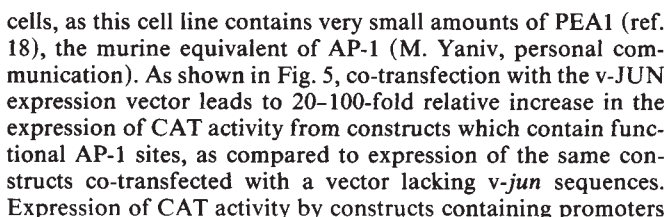
Because the hcJ-1 clone is short (1.5 kb), it cannot contain the complete *c-jun* cDNA. The remaining *c-jun* 5' sequences were isolated as several clones from a HeLa cDNA library, one

of which (hcJ-2) provided the remaining sequence of the c-JUN protein and the 5' untranslated region (see Fig. 2). The combined *c-jun* cDNA is 2.1 kb long, in good agreement with the size predicted from Northern analysis. The ORF of 331 amino acids starts at the first methionine codon at position 412, and predicts a 37K protein with a sequence very similar to the v-JUN protein, except in three regions where the two proteins diverge, possibly due to in-frame deletions that occurred during the genesis of v-jun. The two proteins have 77% sequence identity if these regions are included. The two sequences are almost identical in their putative DNA-binding domains, where only 2 (out of 118) amino acid substitutions can be found.

To characterize the protein products of the v- and c-jun genes, various portions of the ORFs were cloned into the bacterial expression vector pATH2 (see Fig. 3). This vector (constructed by T. J. Koerner and A. Tzagoloff) allows the expression of ORFs fused to the N-terminal portion of the *trpE* gene. After starvation for tryptophan, *Escherichia coli* cells, harbouring pTE-vJ which contains the complete v-jun coding sequence, express an induced protein of M_r 75K (Bp75^{trpE-vjun}). This is the M_r predicted for a fusion of v-JUN to a section corresponding to 37K of the trpE protein. In cells transformed with pTE-cJ-1, which contains the putative DNA-binding domain of human c-jun, the size of the inducible protein is 55K (Bp55^{trpE-cjun}), as predicted from a fusion of 138 amino acids to the trpE protein. None of these proteins is observed in cells transformed with the parental vector, which express an inducible protein of M_r 37K (Bp37^{trpE}) derived from the N-terminal portion of trpE. Taking advantage of the low solubility of trpE-fusion proteins in buffers of high ionic strength¹⁷, we have prepared fractions which are highly enriched in Bp37^{trpE}, Bp75^{trpE-vjun} and Bp55^{trpE-cjun} (Fig. 3, lanes 5–7). These fusion proteins were analysed for their ability to bind to v-JUN anti-peptide antisera by immunoblotting (Fig. 1). Lanes 7 and 9 show silver-stained trpE-c-JUN and trp-c-JUN fusion proteins after SDS-PAGE, and lanes 8 and 10 show their reactivity with one of the anti-JUN antisera thus confirming their identity.

After removal of urea, these proteins (Fig. 3, lanes 5–7) were subjected to DNase I footprinting¹² to test for sequence-specific DNA-binding resembling that of AP-1. As shown in Fig. 4, the fractions that contain Bp75^{trpE-vjun} and Bp55^{trpE-cjun} protect the same regions of the hMT-II_A (lanes 4 and 5) and SV40 (lanes 10 and 11) promoters against DNase I digestion as does AP-1 purified from HeLa cells (lanes 2, 6 and 8). Also like AP-1, these preparations protect the wild-type binding site (TRE) derived from the collagenase promoter (lanes 12–16), but not the inactive point-mutant TREΔ-72 (lanes 17–21). Slight differences can be seen at the borders of the footprints on the collagenase TRE generated by the two trpE-JUN fusion proteins and AP-1. The significance of these differences is not clear but they could be caused by structural alterations conferred by the trpE domain that affect the interaction of the JUN proteins with DNase I. No other footprints were seen even in the absence of competitor DNA, although these DNA probes contain binding sites for other regulatory proteins. The control fraction containing Bp37^{trpE} also did not show any footprinting activity (lanes 3, 9, 14, 19). These experiments indicate that the v-JUN- and c-JUN-trpE fusion proteins are DNA-binding proteins which recognize a core sequence similar to that recognized by AP-1 purified from human cells. Furthermore, as the C-terminal residues of human c-JUN are sufficient to generate an active and specific DNA-binding protein, the DNA-binding domain of the JUN protein can be localized to its C-terminal region, as was originally suggested by the sequence similarity between v-JUN and GCN4 (ref. 9).

A vector pSV-v-jun was constructed to elicit efficient expression of v-JUN in transfected mammalian cells. The effect of this vector on expression of a reporter gene (CAT), driven by several different promoters with or without AP-1 sites, was tested by co-transfection experiments in undifferentiated F9



Methods. A human fibroblast cDNA library constructed in the vector pSP64 (ref. 28) was screened with two synthetic oligodeoxynucleotide probes (5'-CCGCGCTTTTGGACGCCGCAATTC-TGTTTCTCATGCGTTT-3' and 5'-CAGTTTTTCTTCCAACCTGGCAATCCTTTTCCAACCTTCCT-3') derived from the putative DNA-binding domain encoded by *v-jun* (refs 8, 9). Screening of ~80,000 clones yielded a single clone hcJ-1 which gave a strong hybridization signal with both probes. To isolate additional *c-jun* cDNA clones, a HeLa-cell cDNA library was prepared in λ gt11, and $\sim 0.5 \times 10^5$ phage were screened with a nick-translated fragment derived from hcJ-1 to yield a large number of cross-hybridizing clones. One of them, hcJ-2, appeared to contain the largest 5' region and was purified and subcloned into pUC18. Restriction fragments of hcJ-1 and hcJ-2 were subcloned into M13 vectors and their sequence (both strands) determined by the dideoxy-chain-termination method²⁹.

lacking functional AP-1 sites is increased by only 2-4-fold. In experiments presented elsewhere, we show that the CAT enzyme activity expressed by these vectors correlates with the number of correctly initiated transcripts from the TK, collagenase and hMT-II_A promoters^{12,19}. Thus, v-JUN appears to function as a sequence-specific *trans*-activator, which interacts with the same DNA sequences recognized by AP-1 and the trpE-JUN fusion proteins *in vitro*. The reason for the low degree of trans-activation

Fig. 3 Expression of *trpE*-JUN and *trpE*-cJUN fusion proteins in *E. coli*. **a**, Construction of plasmids used for expression. pTE-vJ: the *v-jun* sequence from the *Sac*II site at position 672 to the *Eco*RI site at ~100 base pairs 3' of the stop codon of *v-jun* (ref. 8) was used. After converting the *Eco*RI site to a *Hind*III site and blunt-ending the *Sac*II-site, the 1.1-kb *v-jun* sequence was inserted between the *Sma*I and *Hind*III sites in the polylinker of pATH2, containing the *trpE*-coding region. pTE-cJ-1: the 1.2-kb *Nae*I-*Bam*HI fragment of hc-J-1, containing the complete putative DNA-binding domain of *c-jun*, was cloned between the *Sma*I and *Bam*HI sites in the polylinker of pATH2. The open boxes denote the *trpE* coding region, the hatched boxes represent the coding region of the inserted sequences, and the cross-hatched sequences represent either viral sequences (*v-jun*) or 3' untranslated sequences (*c-jun*). The broken lines between *v-jun* and *c-jun* indicate the highly conserved putative DNA-binding domain. The protein products encoded by the original *trpE* gene and the *trpE*-*jun* fusion genes are indicated on the right. **b**, Expression of Bp75^{*trpE*-*vjun*} and Bp55^{*trpE*-*cjun*} in bacteria. Lanes 1-4: *E. coli* DH5 cells, transformed with the plasmids pATH2, pTE-vJ and pTE-cJ-1, were grown for 24 h at 37 °C in M9 medium supplemented with 0.5% casamino acids, 10 µg ml⁻¹ thiamine and 50 µg ml⁻¹ ampicillin in the presence of L-tryptophan (10 µg ml⁻¹). The cultures were then diluted 100-fold in the same medium and grown in the absence or presence of tryptophan for 12 h. The cells were pelleted, resuspended in SDS-sample buffer, incubated for 10 min at 95 °C, and analysed on 10% SDS-polyacrylamide gels using silver-staining. Minus and plus signs indicate whether expression of the *trpE* genes was repressed or induced. Lanes 5-7, the various *trpE*-derived proteins were enriched by preparing extracts as described¹⁷, except that for induction of the *trpE* and *trpE*-fusion proteins, the protocol described above was used. The insoluble protein fraction was solubilized in 6 M urea, dialysed three times for 45 min at 4 °C against 20 mM HEPES, pH 7.9, 100 mM KCl, 12.5 mM MgCl₂, 0.1 mM EDTA, 1 mM dithiothreitol, 20% glycerol. The bands that migrate at the predicted positions for Bp37^{*trpE*}, Bp75^{*trpE*-*vjun*}, and Bp55^{*trpE*-*cjun*} are marked by a dot. The migration positions of molecular weight standards (lane M) are shown to the left.

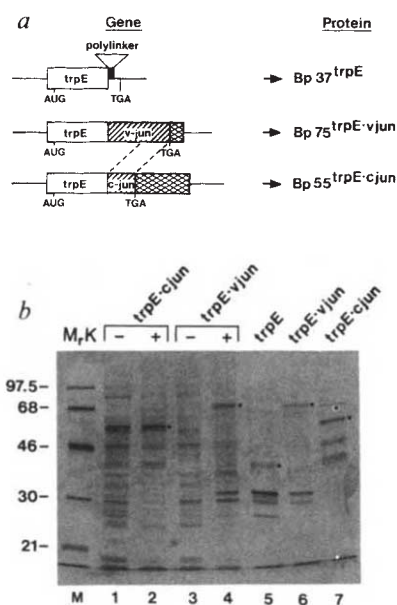
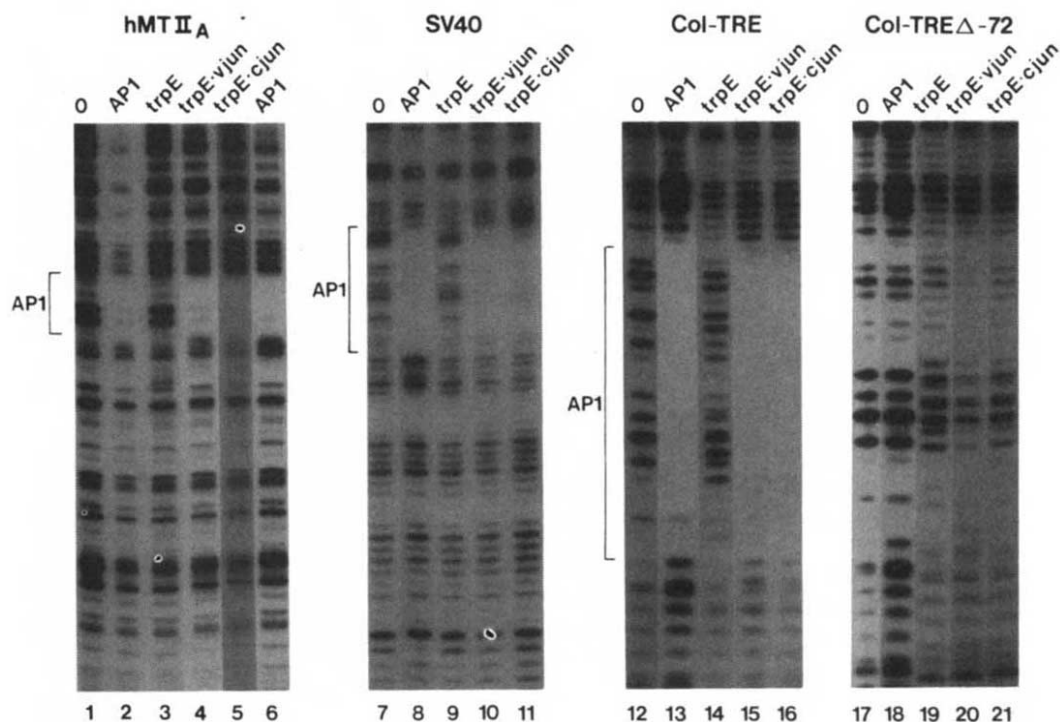


Fig. 4 Footprint analysis of the DNA-binding specificities of AP-1, Bp55^{*trpE*-*cjun*} and Bp75^{*trpE*-*vjun*}. End-labelled hMT-II_A, SV40, TRE collagenase (wild-type) and TREΔ-72 probes were incubated with ~25 ng affinity-purified AP-1 and 1 µg partially purified and renatured Bp37^{*trpE*}, Bp55^{*trpE*-*cjun*}, and Bp75^{*trpE*-*vjun*}, and subjected to DNase I footprinting as previously described¹². The regions protected by AP-1 on each of these probes, with the exception of TREΔ72, are indicated. The hMT-II_A probe was labelled by filling-in of the *Hind*III site at -286 of Δ5' - 286 (ref. 30), and thus represents the anti-sense strand. The SV40 probe was labelled by filling-in of the *Bam*HI site at position 101 of pW2 (P. Chambon, personal communication), and thus represents the sense strand, and the two TRE probes were



labelled by filling-in of the *Eco*RI site at -80 within the HSV-TK promoter to which these TREs were fused¹², and thus represent the sense-strands of TRE col-CAT and TREΔ-72-CAT.

of genes lacking functional AP-1 sites by v-JUN is not clear, but could result from binding of v-JUN to vector sequences resembling AP-1 sites.

Recently it was shown that the C-terminus of v-JUN can functionally replace the GCN4 DNA-binding domain *in vivo*²⁰, but there was no direct demonstration of DNA-binding by the chimaeric protein. After submission of this manuscript, others¹⁶ have also noted the similarity between c-JUN and AP-1 and have found that the sequences of four short peptides derived from AP-1 match sequences present in c-JUN.

Single strongly cross-hybridizing species in digests of human DNA indicates the presence of a solitary *c-jun* gene which gives rise to a simple pattern of transcripts that differ in their 3'-untranslated regions probably as a result of alternative poly-

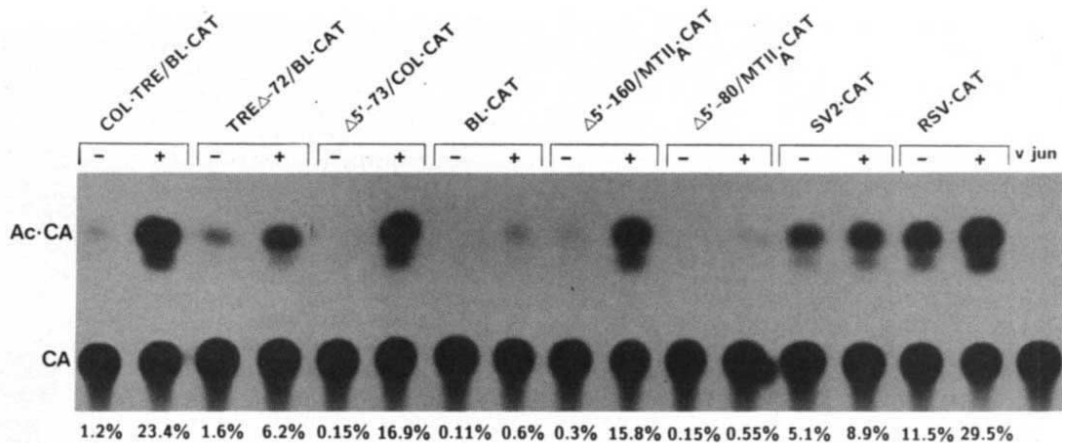
adenylation sites. Thus we expect HeLa cells to express a single *c-jun* gene-encoded product of predicted *M_r* 37K, which is lower than that of AP-1 (44-45K). But as AP-1 is a phosphoprotein (F. Mercurio and W.J.B., unpublished results), it should be retarded on SDS-PAGE (ref. 21). Unless the c-JUN protein has the same *M_r* as AP-1, it should be detected by immunoblotting as a distinct species copurifying with AP-1 as it binds to the same DNA sequence. But the weak cross-hybridizing genomic DNA fragments suggest that c-JUN is a member of a family of DNA-binding proteins that may recognize related, but not identical DNA sequences. One of these, JUN-B, was recently isolated as the product of a serum-inducible gene from mouse fibroblasts²².

Consistent with the sequence similarities between v-JUN and

Fig. 5 Sequence-specific transactivation by *v-jun*. To demonstrate transactivation of transcription by *v-jun*, the various constructs (1 µg DNA per plate) indicated in the upper panel were transfected into F9 cells together with 20 µg pSV-*v-jun* (+) or pUC-SV(-). After 34 h the CAT enzyme activity expressed by the different vectors was determined as described¹².

Col-TRE/BL-CAT and TREΔ-72/BL-CAT contain wild type and mutant TREs in front of the HSV-TK promoter fused to the CAT gene¹². Δ5'-73/Col-CAT

contains a 5'-deleted collagenase promoter (containing an AP-1 site) fused to the CAT gene³¹. BL-CAT contains a promoterless CAT gene¹², Δ5'-160/MTII_A-CAT (containing the AP-1 site) and Δ5'-80/MTII_A-CAT (lacking the AP-1 site) are 5' deletion mutants of the hMTII_A promoter fused to the CAT gene^{19,30}, whereas SV2CAT and RSV-CAT contain the SV40 and RSV promoters fused to the CAT gene¹². The *v-JUN* expression vector pSV-*v-jun* was constructed by inserting the *Sac*II-*Eco*RI fragment, which has the complete *v-jun* sequence (see Fig. 2) downstream to the SV40 promoter region, contained within a *Kpn*I/*Hind*III fragment (positions 294 to 5,171) subcloned into pUC19 (pUC-SV). An 847-base pair *Bgl*III-*Bam*HI fragment derived from SV40, containing the required RNA splicing and polyadenylation signals, was inserted downstream to the *v-jun* sequence. The conversion of [¹⁴C]chloramphenicol (CA) to the monoacetylated form (AC-CA) is indicated at the bottom panel. The lane marked *v-jun* indicates an extract of cells transfected with pSV-*v-jun* in the absence of an indicator plasmid.



GCN4 (ref. 9) and with the result of the GCN4-*jun* exchange experiment²⁰, we find that the carboxy-terminal region of c-JUN contains its DNA-binding domain. Yet a smaller trpE-c-JUN fusion protein, containing the last 99 C-terminal amino acids of c-JUN is severely impaired in its DNA-binding activity (unpublished results). Interestingly, a GCN4-*v-jun* recombinant containing the equivalent region of *v-jun* (LJG-2), was impaired in its GCN4 complementation activity²⁰, suggesting that the DNA-binding domain of the JUN proteins could encompass more than just the GCN4-homologous region. Although the similarity between the two JUN sequences and GCN4 is limited to the C-terminal region, the similarity between the *v-JUN* and the c-JUN sequences extends to their N-termini, but is not maintained throughout the entire sequence as it diverges over three short regions. This divergence could be due to evolutionary differences between chicken and man affecting functionally unimportant regions of the molecule, or to accumulation of mutations required for generation of a constitutively active *v-jun* oncogene from the *c-jun* proto-oncogene.

In addition to functioning as a basic transcription factor^{13,14}, AP-1 is an enhancer-binding factor that mediates induction by phorbol ester tumour promoters^{12,15}. Phorbol esters such as TPA (12-*O*-tetradecanoyl phorbol 13-acetate) function by activating protein kinase C (ref. 22), so it has been proposed that AP-1 acts at the nuclear end of a complex signal transduction pathway starting at the cell surface and ultimately responsible for increased expression of specific genes. Other participants in this pathway are proteins controlling the production of diacylglycerols, the physiological activators of protein kinase C (ref. 23). The c-Ha-ras oncogene can activate the polyoma enhancer in a manner similar to TPA (ref. 24) an effect mediated by the factor PEA-1 (B. Wasylyk, personal communication). Therefore, it is likely that c-Ha-ras participates in the same pathway as *c-jun* but at an earlier step mediated by its effect on diacylglycerol metabolism²⁵.

Another member of the nuclear-acting oncogenes is *v-erb-A*, whose cellular counterparts code for thyroid hormone receptors^{6,7}; the mechanism by which these proteins regulate transcription is not well understood because thus far they have not been shown to directly stimulate it. On the other hand, AP-1 and *v-JUN* stimulate transcription *in vitro*^{13,14} and *in vivo* (this work). Although AP-1 could be the first reported transcription factor to be encoded by a proto-oncogene, a factor that mediates

the action of tumour promoters could be expected to have oncogenic potential. Our findings strongly suggest that *v-JUN* acts by increasing the expression of specific genes to establish the transformed phenotype. This may be accomplished by loss of proper regulation or by production of a structurally altered protein from a mutated gene. The oncogene *v-jun* may differ from its cellular counterpart, *c-jun*, by no longer requiring the activation of protein kinase C to maintain maximal activity—that is, possible mutations in the regulatory domain could allow constitutive DNA binding and transcriptional activation. In the future it will be important to identify the *c-jun*/AP-1 activated genes that are directly responsible for neoplastic transformation, and to determine whether other transcription factors involved in signal transduction, such as AP-2 (ref. 26), are also involved in oncogenesis.

We thank Drs Tim Bos and Peter Vogt for the generous gift of anti-JUN antisera before publication, and the *v-jun* clone; Dr Robert Chiu for initial characterization of *c-jun* messenger RNA using the oligonucleotide probes; Dr Joe Lipsick for a suggestion to use trpE-expression vectors and oligonucleotide synthesis; Pamela Bucher and Susan Smith for preparation of the manuscript and Jennifer Meek for artwork. This work was supported by grants from the National Institutes of Health, Environmental Protection Agency, Department of Energy, American Cancer Society and Council for Tobacco Research to M.K. and T.H., P.A. and W.J.B. are supported by postdoctoral fellowships from the Deutscher Akademischer Austauschdienst and the Life Science Foundation respectively, and S.O. holds a predoctoral NIH training grant.

Received 2 November 1987; accepted 6 January 1988.

1. Bishop, J. M. *Science* **225**, 305-311 (1987).
2. Weinberg, R. A. *Science* **230**, 770-776 (1985).
3. Heldin, C. H. & Westermark, B. *Cell* **37**, 9-20 (1984).
4. Downward, J. *et al. Nature* **307**, 521-527 (1984).
5. Sherr, C. J. *et al. Cell* **41**, 665-676 (1985).
6. Sap, J. *et al. Nature* **324**, 635-640 (1986).
7. Weinberger, C. *et al. Nature* **324**, 641-646 (1986).
8. Maki, Y. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 2848-2852 (1987).
9. Vogt, P. K., Bos, T. J. & Doolittle, R. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3316-3319 (1987).
10. Hope, I. A. & Struhl, K. *Cell* **46**, 885-894 (1986).
11. Hope, I. A. & Struhl, K. *Cell* **43**, 177-188 (1985).
12. Angel, P. *et al. Cell* **49**, 729-739 (1987).
13. Lee, W., Mitchell, P. & Tjian, R. *Cell* **49**, 741-752 (1987).
14. Lee, W., Haslinger, A., Karin, M. & Tjian, R. *Nature* **325**, 368-372 (1987).
15. Chiu, R., Imagawa, M., Imbra, R. J., Bockoven, J. R. & Karin, M. *Nature* **329**, 648-651 (1987).
16. Bohmann, D. *et al. Science* **238**, 1386-1392 (1987).

17. Klempnauer, K.-H. & Sippel, A. E. *EMBO J.* **6**, 2719-2725 (1987).
18. Krysky, M. H., Piette, J. & Yaniv, M. *Nature* **328**, 254-256 (1987).
19. Karin, M. *et al. Molec. cell. Biol.* **7**, 606-613 (1987).
20. Struhl, K. *Cell* **50**, 841-846 (1987).
21. Barber, J. R. & Verma, I. M. *Molec. cell. Biol.* **7**, 2201-2211 (1987).
22. Ryder, K., Lau, L. F. & Nathans, D. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
23. Nishizuka, Y. *Nature* **308**, 693-698 (1984).
24. Waslyk, C., Imler, J. L., Perez-Mutul, J. & Waslyk, B. *Cell* **48**, 525-534 (1987).
25. Fleischman, L. F., Chahwala, S. B. & Cantley, L. *Science* **231**, 407-410 (1986).
26. Imagawa, M., Chiu, R. & Karin, M. *Cell* **51**, 251-260 (1987).
27. Towbin, H., Starhelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
28. Whitham, S. E. *et al. Biochem. J.* **240**, 913-916 (1986).
29. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
30. Karin, M. *et al. Nature* **308**, 513-519 (1984).
31. Angel, P. *et al. Molec. cell. Biol.* **7**, 2256-2266 (1987).

Translational activation of the *lck* proto-oncogene

Jamey D. Marth*§, Robert W. Overell^{||},
Kathryn E. Meier*‡, Edwin G. Krebs*‡
& Roger M. Perlmutter*†§

* Howard Hughes Medical Institute, † Departments of Biochemistry,
‡ Pharmacology and § Medicine, University of Washington, Seattle,
Washington 98195, USA

^{||} Immunex Corporation, 51 University Street, Seattle,
Washington 98101, USA

The *lck* proto-oncogene, a member of the *src* gene family, encodes a lymphocyte-specific protein tyrosine kinase ($p56^{lck}$) that is implicated in the pathogenesis of lymphoid neoplasia¹⁻⁴. We report here that 5' *lck* sequence elements, containing AUG codons, significantly reduce the *in vivo* efficiency of $p56^{lck}$ translation from the normal messenger RNA. This result provides a quantitative explanation for the overexpression of $p56^{lck}$ in two retrovirally-induced murine lymphomas that express abnormal *lck* transcripts from which the physiological 5' translational control region has been deleted and non AUG-containing viral sequences substituted. In contrast to other mammalian genes, most proto-oncogenes contain AUG codons 5' to the authentic initiation codon⁵, suggesting that cells can regulate translational start sites in these mRNAs. Abrogation of translational control may be a general mechanism of proto-oncogene activation in malignancy.

In the murine lymphoma LSTRA, one allele of the *lck* locus is rearranged due to the insertion of Moloney murine leukaemia virus (MoMuLV) approximately 1 kilobase (kb) 5' of the normal transcriptional start site (refs 1, 2 and 6, and manuscript in preparation). As a result, *lck* transcripts from the rearranged allele initiate within the MoMuLV long terminal repeat (LTR), and contain 204 bases of retroviral sequence spliced into an acceptor site five bases 5' to the initiation codon for $p56^{lck}$ (ref. 2, and manuscript in preparation). In contrast, the normal *lck* mRNA includes 177 bases of 5'-untranslated region that are deleted from the LTR-*lck* fusion transcript (Fig. 1a). Both the normal and the fusion transcripts encode identical $p56^{lck}$ products¹⁻³.

Using a probe derived from the *lck* 5' untranslated region, YAC-1 cells (a MoMuLV-transformed lymphoma that does not have an *lck* gene rearrangement¹) contain 1.7-fold more *lck* mRNA than do LSTRA cells (ref. 7, Fig. 1b). These data indicate that transcripts from the unrearranged *lck* allele in LSTRA cells accumulate at a relatively normal rate. But when a full-length *lck* complementary DNA probe is used on mRNA blots, the abundance of *lck* mRNA is 7-fold greater in LSTRA cells than in YAC-1 (Fig. 1b). Thus the measured increase in *lck* mRNA in LSTRA cells reflects the accumulation of the MoMuLV LTR-*lck* fusion transcript.

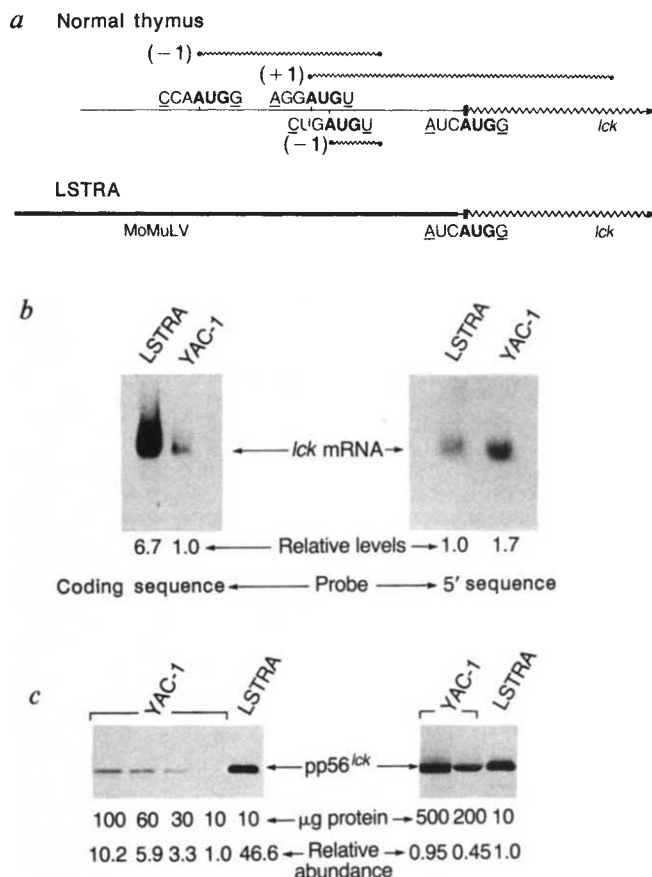


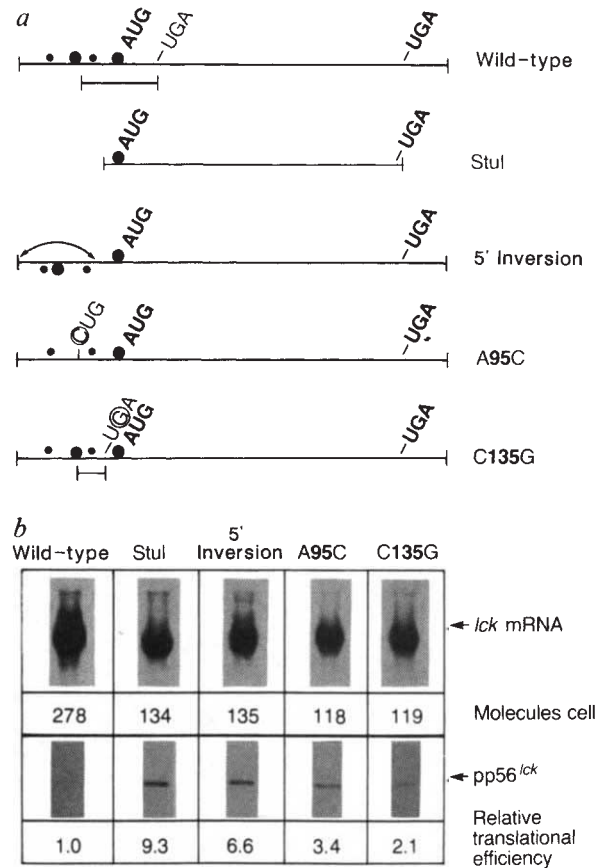
Fig. 1 MoMuLV-mediated truncation of 5' *lck* sequences correlates with disproportionately high levels of the $p56^{lck}$ protein tyrosine kinase. **a**, Comparison of 5' *lck* sequences expressed in either normal mouse thymus or LSTRA *lck* mRNAs. AUGs for $p56^{lck}$ and those residing 5' are presented with their respective nucleotide sequence contexts. Studies by Kozak^{10,11} demonstrate that optimal translational initiation occurs when purines are found at the -3 position (underlined) relative to the A of the AUG. The +4 position (underlined) is optimally G, but may be other nucleotides when a purine is present at -3 (ref. 11). Open reading frames beginning at AUGs are denoted by jagged lines. An erroneous single-base insertion (position 65) in the published NT18 *lck* cDNA sequence has been corrected¹. **b**, Using the 1.75 kb NT18 *lck* cDNA as a probe¹, poly(A)⁺ RNA blot analysis and solution hybridization revealed a 6.7-fold increase in *lck* mRNA in LSTRA as compared to YAC-1 cells⁷ (left). When a 5' *lck* fragment (obtained by *EcoRI/StuI* digestion of the NT18 cDNA) is used as a probe, a 1.7-fold increase in *lck* mRNA, quantitated by laser densitometry, is measured in the YAC-1 cell line (right). **c**, Immunoblot analyses of LSTRA or YAC-1 membrane proteins using affinity-purified $p56^{lck}$ antibodies⁷. The left and right-hand panels represent two separate experiments.

Methods. For each lane, the amount of Triton-soluble membrane protein assayed by immunoblotting⁷ is indicated as is the amount of $p56^{lck}$ protein quantitated using an LKB laser densitometer. In the left-hand panel, these values were normalized to the level of $p56^{lck}$ measured when 10 μg of YAC-1 membrane protein was assayed. In the right-hand panel, the values were normalized to levels measured using 10 μg of LSTRA membrane protein.

The abundance of $p56^{lck}$ protein in LSTRA and YAC-1 cells was determined by immunoblotting with an *lck*-specific antiserum⁷. When the amount of YAC-1 membrane protein assayed is adjusted to reflect the degree of *lck* mRNA overexpression in LSTRA cells, $p56^{lck}$ in LSTRA membranes is still in vast excess (Fig. 1c). These titrations reveal that LSTRA cells contain 50-fold more $p56^{lck}$ protein than do YAC-1 cells (Fig. 1c). Extending these analyses to normal murine thymocytes or other

Fig. 2 The 5' *lck* mRNA sequence contains elements that reduce the translational efficiency of $p56^{lck}$ *in vivo*. **a**, Schematic representations of *lck* cDNA constructs. The AUG and UGA signals for $p56^{lck}$ are shown in bold type. AUGs are denoted with darkened circles. The larger circles represent AUGs in good ribosomal sequence contexts expected to permit efficient translational initiation. The open reading frame starting from the second 5' AUG is indicated in the wild-type construct and in the C135G construct where it was deliberately truncated. **b**, Results from analyses of polyclonal cell lines transfected with the above *lck* cDNAs expressed from the mouse MT-I promoter²⁷. *lck* mRNA was visualized by Northern blotting and quantitated by solution hybridization techniques as previously⁷. The amount of membrane protein used in SDS-PAGE and immunoblotting analyses⁷ was then adjusted such that each assay contained $p56^{lck}$ levels derived from an identical amount of *lck* mRNA. The relative translational efficiency of each *lck* mRNA transcript is presented at the bottom of the figure normalized to the level of $p56^{lck}$ measured in the wild-type cell line.

Methods. To produce the A95C or C135G *lck* cDNA fragments, site-directed mutagenesis was accomplished as previously⁴ using oligomers of sequences: 5'CATCAGACAGCCTAGAGCCC 3' or 5'GAGAGGCCTCAAGGAGCTGC 3' which span nucleotides 85–104 or 126–146 of the NT18 *lck* cDNA clone¹. Sequencing analyses confirmed that the specific mutation was incorporated into *lck* sequences before reconstruction of full length cDNA clones and insertion into the *EcoRI* site of the pNUT vector²⁷ through blunt-ended ligations. The 5' Inversion construct was created by a blunt-ended ligation of the 5' *EcoRI*-*StuI* 138 base NT18 fragment back into an *lck* cDNA (*StuI* fragment) vector. These procedures did not result in the generation of 5' AUGs. BALB/c 3T3 cells were transfected as previously⁴ and polyclonal cell lines containing over 200 distinct clones were selected by methotrexate resistance (0.1 μ M). This figure was generated using samples derived from a single representative determination analysed for both *lck* mRNA and $p56^{lck}$ protein simultaneously. The NT18 *lck* cDNA clone¹ contains 17 base pairs (bp) (without AUGs) adjacent to the 5' *EcoRI* linker that are derived from a cloning artefact, 145 bp of 5' untranslated region, and the complete coding region for $p56^{lck}$. Similar results were observed when the experiment was repeated using a transient transfection protocol (data not shown).



lymphoma cell lines yielded similar results: LSTRA membranes contain 7 to 10-fold more $p56^{lck}$ protein than would be predicted from the measured increment in *lck* mRNA expression (data not shown). Thus the normal *lck* mRNA appears to be less efficiently translated than is the viral LTR-*lck* fusion transcript produced in LSTRA cells.

The 5' untranslated region of the normal *lck* mRNA contains three AUGs, all of which are out-of-frame with respect to the initiation AUG of $p56^{lck}$ (Fig. 1a). However, the LTR-*lck* fusion transcript does not contain 5' AUGs. Previous studies of ribosome-mRNA recognition indicate that 40S ribosomal subunits scan mRNA sequences from their 5' caps^{8–10}. Translation begins when an AUG codon is encountered in an appropriate sequence context, typically a purine residue at position -3 with respect to the A of the AUG^{11–12}. Of the three 5' AUGs in the *lck* mRNA, the second lies in a sequence context especially suitable for recognition by the 40S mammalian ribosomal subunit (Fig. 1a). Translation from this start site would continue in the +1 reading frame for 49 codons, terminating 80 bases 3' to the AUG for $p56^{lck}$ (Fig. 1a). Based on previous studies^{13–15}, this distance is probably too large to permit translational re-initiation at the *bona fide* AUG for $p56^{lck}$.

To test the efficacy of the *lck* 5' sequence in suppressing translation, $p56^{lck}$ expression constructs bearing modifications in this region were assayed for translational efficiency. The NT18 *lck* cDNA clone¹ was used as the wild-type construct (Fig. 2a). Another construct deleted 138 5' bases including all three 5' AUGs. To compare translation from *lck* mRNA species whose 5' sequence length is identical to the wild-type construct, a third construct was generated in which 138 5' bases were inverted, eliminating all 5' AUGs. Lastly, as it seemed likely that the second 5' AUG in the *lck* mRNA could mask $p56^{lck}$ translation, site-directed mutagenesis was used to convert this AUG to a CUG, or to generate a termination codon that would halt transla-

tion 27 bases 5' to the AUG codon of $p56^{lck}$. The wild-type and mutant *lck* expression constructs were transfected into BALB/c 3T3 fibroblast cells and stable polyclonal *lck*-expressing cell lines were obtained (Fig. 2). Measurements of *lck* mRNA abundance and $p56^{lck}$ production allowed an assessment of the translational efficiency of each *lck* construct.

Truncation of 138 bases of 5' sequence to leave 29 bases of 5' untranslated region (24 from the *lck* gene and five from the metallothionein I gene) results in a 9- to 10-fold increase in translational efficiency compared to the wild-type construct (Fig. 2b). Similarly, inversion of the 138 5' bases generates a transcript of wild-type length that translates $p56^{lck}$ 6- to 7-fold more efficiently than the wild-type construct. These results focus attention on sequence elements within the 5' portion of the *lck* mRNA that mask $p56^{lck}$ translation, for example, the 5' AUG triplets.

Elimination of the second 5' AUG results in a 3- to 4-fold increase in the translational efficiency of $p56^{lck}$ from the *lck* mRNA (Fig. 2b). Thus ~50% of the increase in the translational efficiency of $p56^{lck}$ achieved by inverting the *lck* 5' region appears to be due to ribosomal recognition of the second 5' AUG. Premature termination of the open reading frame proceeding from this start site to permit re-initiation at the AUG for $p56^{lck}$ (C135G construct) yields a 2-fold increase in translation. These results are in accord with studies of translational re-initiation efficiency in mammalian cells performed using other expression systems^{13–15}. As both the inversion and truncation constructs direct the synthesis of $p56^{lck}$ better than does the A95C construct, other 5' elements, including the two other 5' AUGs, probably contribute to the inefficient translation of $p56^{lck}$ from the normal *lck* mRNA.

Considerable evidence implicates $p56^{lck}$ in the pathogenesis of murine and human neoplastic disease^{1,2,4,16}. The *lck* gene is a target for retroviral insertion in at least two murine lymphoma cell lines, and is positioned at a site of frequent chromosomal

abnormalities (1p32-35) in human malignant disease³. Moreover, alteration of the carboxy-terminal phosphorylation site of p56^{lck} (Tyr 505 to Phe 505) generates a protein that can fully transform NIH 3T3 cells⁴. Phosphopeptide mapping studies suggest that Tyr 505 of p56^{lck} is relatively hypophosphorylated in LSTRA cells, implying that although the primary structure of p56^{lck} in LSTRA cells is normal, some p56^{lck} molecules may nevertheless be activated^{4,17,18}. Consistent with this view, the LSTRA cell line contains greatly increased levels of phosphotyrosine^{19,20}.

Our results provide a quantitative explanation for the 50-fold overexpression of p56^{lck} in LSTRA cells. First, retroviral promoter insertion generates a 6- to 7-fold increase in *lck* mRNA levels. Second, as a result of RNA splicing mechanisms, 5' *lck* sequence elements (including AUGs) that normally modulate p56^{lck} translational initiation are removed. Thus the MoMuLV LTR-*lck* mRNA transcribed from the rearranged allele, and lacking 5' AUGs, may be translated 6- to 9-fold more efficiently. Under certain circumstances the phosphorylation state of p56^{lck} may change when overexpressed, thereby increasing enzymatic activity and unmasking its transforming potential⁴.

Translational activation of proto-oncogenes through the deletion of 5' AUG elements probably contributes to oncoprotein overexpression in other transformed cells. The only other known *lck* gene rearrangement, in the murine thymoma Thy-19, also results from a 5' MoMuLV insertion, and yields a viral LTR-*lck* fusion transcript in which the normal 5' *lck* sequences are eliminated, and p56^{lck} production is increased, exactly as in LSTRA cells^{6,21}. Translational activation may also explain the increased levels of platelet-derived growth-factor B chain produced by some transforming versions of the *sis* oncogene in which the normal 5' untranslated region (containing out-of-frame AUGs) is truncated^{22,23}. Although fewer than 10% of 700 surveyed mammalian genes have AUGs 5' to the authentic initiation codon, over 65% of proto-oncogene transcripts contain these elements⁵. For example, at least five of the seven known members of the *src* gene family (*fgr*, *fyn*, *hck*, *lck* and *src*) contain multiple 5' AUGs, virtually all of which are in sequence contexts appropriate for efficient ribosome recognition.

These observations suggest two non-exclusive hypotheses: first, that 5' AUGs assist in preventing the inappropriate overexpression of gene products that can regulate cell growth; and second, that translation of these sequences can be regulated by controlling the extent of recognition of AUGs 5' of the authentic initiation codon. Indeed, this type of regulation has been documented in the control of translation of the *GCN4* locus in yeast²⁴⁻²⁶. In light of the large number of proto-oncogene transcripts that contain 5' AUGs, elimination of these elements, as in the generation of retroviral LTR-*lck* fusion transcripts in both LSTRA and Thy-19 lymphomas, may be a common mechanism of oncogene activation.

Received 1 December 1987; accepted 26 January 1988.

- Marth, J. D., Peet, R., Krebs, E. G. & Perlmutter, R. M. *Cell* **43**, 393-404 (1985).
- Voronova, A. F. & Sefton, B. M. *Nature* **319**, 682-685 (1986).
- Marth, J. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 7400-7404 (1986).
- Marth, J. D. *et al. Molec. cell. Biol.* **8**, 540-550 (1988).
- Kozak, M. *Nucleic Acids Res.* **15**, 8125-8148 (1987).
- Voronova, A. F., Adler, H. T. & Sefton, B. M. *Molec. cell. Biol.* **7**, 4407-4413 (1987).
- Marth, J. D. *et al. EMBO J.* **9**, 2727-2734 (1987).
- Kozak, M. & Shatkin, A. J. *J. biol. Chem.* **253**, 6568-6577 (1978).
- Kozak, M. *Cell* **22**, 459-467 (1980).
- Kozak, M. *Cell* **34**, 971-979 (1983).
- Kozak, M. *Nature* **308**, 241-246 (1984).
- Kozak, M. *Cell* **44**, 283-292 (1986).
- Peabody, D. S., Subramani, S. & Berg, P. *Molec. cell. Biol.* **6**, 2704-2711 (1986).
- Peabody, D. S. & Berg, P. *Molec. cell. Biol.* **6**, 2695-2703 (1986).
- Kozak, M. *Molec. cell. Biol.* **7**, 3438-3445 (1987).
- Veillette, A., Foss, F. M., Sausville, E. A., Bolen, J. B. & Rosen, N. *Oncogene Res.* **1**, 357-374 (1987).
- Cooper, J. A. & King, C. S. *Mol. cell. Biol.* **6**, 4467-4477 (1986).
- Cooper, J. A. in *Peptides and Protein Phosphorylation* (eds Kemp, B. & Alewood, P. F.), CRC, Boca Raton, Florida, in the press.
- Gacon G. *et al. EMBO J.* **1**, 1579-1582 (1982).

- Casnellie, J. E., Harrison, M. L., Hellstrom, K. E. & Krebs, E. G. *J. biol. Chem.* **258**, 10738-10742 (1983).
- Voronova, A. F. *et al. Molec. cell. Biol.* **4**, 2705-2713 (1984).
- Collins, T., Ginsburg, D., Boss, J. M., Orkin, S. H. & Pober, J. S. *Nature* **316**, 748-750 (1985).
- Ratner, L., Thielan, B. & Collins, T. *Nucleic Acids Res.* **15**, 6017-6036 (1987).
- Mueller, P. P. & Hinnebusch, A. G. *Cell* **45**, 201-207 (1986).
- Tzamaras, D., Alexandraki, D. & Thireos, G. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4849-4853 (1986).
- Mueller, P. P., Harashima, S. & Hinnebusch, A. G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2863-2867 (1987).
- Palmiter, R. D. *et al. Cell* **50**, 435-443 (1987).

Hybrid penicillin-binding proteins in penicillin-resistant strains of *Neisseria gonorrhoeae*

Brian G. Spratt

Microbial Genetics Group, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK

Benzylpenicillin has been used extensively for ~40 years in the treatment of gonorrhoea. The intense selective pressures resulting from the continual exposure of *Neisseria gonorrhoeae* to penicillin have resulted in the emergence of resistant strains that produce altered forms of penicillin-binding proteins (PBPs) with decreased affinity for the antibiotic¹⁻³. A comparison of the sequences of the PBP-2 genes from penicillin-sensitive and penicillin-resistant strains, suggests that penicillin-resistant forms of PBP 2 may have arisen both by amino-acid substitutions and insertions, and by the exchange of a region encoding part of the penicillin-sensitive transpeptidase domain with the homologous region from a closely related species.

The penicillin sensitivity of clinical isolation of *N. gonorrhoeae* has decreased over the past two decades and strains with minimal inhibitory concentrations (MICs) of $>1 \mu\text{g ml}^{-1}$ are now encountered in most countries⁴. Penicillin resistance in non- β -lactamase-producing gonococci (chromosomally-mediated resistant *N. gonorrhoeae*; CMRNG) is caused by the production of altered forms of PBPs 1 and 2 that have decreased affinity for the antibiotic, and by mutations that reduce the permeability of the outer membrane.

Penicillin resistance can be transferred from a CMRNG strain into a penicillin-sensitive strain in a stepwise way by transformation¹⁻⁵. The first level of resistance introduced into the penicillin-sensitive strain, FA19, by transformation with DNA isolated from the CMRNG strain CDC77-124615, results from the introduction of the altered PBP-2 gene (*penA*) of the latter strain¹. The *penA* gene of *N. gonorrhoeae* was therefore cloned by constructing a gene library of DNA fragments of strain CDC77-124615 in *Escherichia coli*, and identifying recombinant plasmids that could transform the gonococcal strain FA19 to the first level of penicillin resistance. The nucleotide sequences of the *penA* gene of two penicillin-sensitive strains and three CMRNG strains are compared in Fig. 1.

The deduced relative molecular mass of PBP 2 is 63,650 which agrees with the previously reported value of ~60,000 (refs 1, 2). Expression of the *penA* gene in *E. coli* resulted in the appearance of a novel PBP with the same electrophoretic mobility as PBP 2 of *N. gonorrhoeae* (data not shown). Comparison of the amino-acid sequence of PBP 2 of *N. gonorrhoeae* with the sequences of the four high relative-molecular-mass PBPs of *E. coli*⁶⁻⁸ (Fig. 2) showed that the gonococcal PBP is the homologue of *E. coli* PBP 3 (40% amino-acid sequence similarity) and can therefore be assigned a role as the penicillin-sensitive enzyme involved in peptidoglycan synthesis at cell division⁹. PBP 3 of *E. coli* is believed to be bifunctional, with an amino-terminal penicillin-insensitive transglycosylase domain, and a carboxy-terminal penicillin-sensitive transpeptidase domain¹⁰⁻¹². The transpeptidase domain of the gonococcal

	1	20	
	MetLeuIleLysSerGluTyrLysProArgMetLeuProLysGluGluGlnValLysLysProMetThrSerAsnGlyArgIleSerPheValLeu		
	GGATAATAATAACGAGAAGTAAAAATGTTGATTAAAGCGAATAAAGCCCGGATGTGCCCAAGAGAGCAGGTCAAAAAGCCGATGACCAAGTACCGGATTAGCTTCGTCTCTG	120	
	40	60	
	MetAlaMetAlaValLeuPheAlaCysLeuIleAlaArgGlyLeuTyrLeuGlnThrValThrTyrAsnPheLeuLysGluGlnGlyAspAsnArgIleValArgThrGlnAlaLeuPro		
	ATGGCAATGGCGGTCTTGTGTTGCTGCTGATTGCCCGGGCTGTATCTGCAGACGGTAACTGATACTTTTAAAGAACAGGGCGCAACCGGATTGTGGGACTCAAGCATTGCGG	240	
	80	100	
	AlaThrArgGlyThrValSerAspArgAsnGlyAlaValLeuAlaLeuSerAlaProThrGluSerLeuPheAlaValProLysAspMetLysGluMetProSerAlaAlaGlnLeuGlu		
	GCTACACGGGTACGGTTTCGACCGGAACGGTGGCGTTTGGCGTTGAGCGCGCCGACGGAGTCCCTGTTTGGCGTGCCTAAGATATGAAGAAATGCGCGCCAATTGGAA	360	
	120	140	
	ArgLeuSerGluLeuValAspValProValAspValLeuArgAsnLysLeuGluGlnLysGlyLysSerPheIleTrpIleLysArgGlnLeuAspProLysValAlaGluGluValLys		
	CGCTGTCCGAGCTTGTGATGTGCGGTGCGATGTTTGGAGAACAACTCGAACAGAAAGCAAGTCGTTTATTGGATCAAGCGGCAGCTCGATCCCAAGTTGCCGAAGAGGTCAAA	480	
	160	180	
	AlaLeuGlyLeuGluAsnPheGluPheGluLysGluLeuLysArgHisTyrProMetGlyAsnLeuPheAlaHisValIleGlyPheThrAspPheAspGlyLysGlyGlnGluGlyLeu		
	GCCTTGGGTTTGGAAAATTTGTATTGAAAAAGAAATTAACGCCATTACCCGATGGGCAACCTGTTTGCACACGTCATCGGATTACCGATATTGACGGCAAGGTGAGGAGTTTG	600	
	200	220	
	GluLeuSerLeuGluAspSerLeuTyrGlyGluAspGlyAlaGluValValLeuArgAspArgGlnGlyAsnIleValAspSerLeuAspSerProArgAsnLysAlaProGlnAsnGly		
	GAACTTTCGCTTGAAGACAGCCTGTATGGCGAAGACGGCGCGGAAGTTGTTTGGCGGACCGGCAAGGCAATATTGTGGACAGCTTGGACTCCCCGCGCAATAAAGCACCGCAAAACGGC	720	
	240	260	
	LysAspIleIleLeuSerLeuAspGlnArgIleGlnThrLeuAlaTyrGluGluLeuAsnLysAlaValGluTyrHisGlnAlaLysAlaGlyThrValValValLeuAspAlaArgTh		
	AAAGACATCATCTTTCCTTCGATCAGAGATTACAGACCTTGGCTATGAAGAGTTGAACAGGCGGTGAATACCATCAGGCAAAAGCCGGAACGGTGGTGGTTTGGATGCCCGCAC	840	
	280	300	
	GlyGluIleLeuAlaLeuAlaAsnThrProAlaTyrAspProAsnArgProGlyArgAlaAspSerGluGlnArgArgAsnArgAlaValThrAspMetIleGluProGlySerAlaIle		
	GGGGAATCCCTCGCTTGGCCAAATACGCCCGCTACGATCCCAACAGACCCGGCGGGCAGACAGCGAAGCAGCGGCGCAACCGTGCCTAACCAGATATGATCGAACCTGGTTCCGGCAATC	960	
	320	340	
	LysProPheValIleAlaLysAlaLeuAspAlaGlyLysThrAspLeuAsnGluArgLeuAsnThrGlnProTyrLysIleGlyProSerProValArg AspThrHisValTyrPro		
A)	AAACCGTTCGTGATTGCGAAGGCATTGGATGCGGGCAAAACCGATTGAACGAACGGCTGAATACGCAGCCTTATAAAATCGGACCGTCTCCCGTGGCGC...ATACCATGTTTACCCC	1080	
B)	-----		
C)	-----	Asp	
D)	-----	ACG	
E)	-----	Asp	
	360	380	
	SerLeuAspValArgGlyIleMetGlnLysSerSerAsnValGlyThrSerLysLeuSerAlaArgPheGlyAlaGluGluMetTyrAspPheTyrHisGluLeuGlyIleGlyValArg		
	TCCTTGGATGTGCGCGCATTTATGCAGAAATCGTCCAACGTGCGGCACAAGCAAACTGCTCGCGGTTTCGGCGCGCAAGAAATGTATGACTTCTATCATGAATTGGGCATCGGTGTCGT	1200	
	400	420	
	MetHisSerGlyPheProGlyGluThrAlaGlyLeuLeuArgAsnTrpArgArgTrpArgProIleGluGlnAlaThrMetSerPheGlyTyrGlyLeuGlnLeuSerLeuLeuGlnLeu		
A)	ATGCACTCGGCTTTCGGGGGAACTCAGGTTTGTGAGAAATTTGGCCGAGTGGCGGCCATCGAACAGGCGACGATGCTTTTCGGTTACGGTCTGCAATTGAGCCTGCTGCAATTG	1320	
B)	-----		
C)	-----	C	
D)	-----	C	
E)	-----	C	
	440	460	
	AlaArgAlaTyrThrAlaLeuThrHisAspGlyValLeuLeuProLeuSerPheGluLysGlnAlaValAlaProGlnGlyLysArgIlePheLysGluSerThrAlaArgGluValArg		
	GGCGCGCTATACCGCACTGACGCGACGAGCGGTTTGTCTGCGCTCAGCTTTGAGAGAGCAGGCGGTGCGCGCAAGGCAACGCATATTCAAGAATCGACCGCGCGAGGTACGC	1440	
	480	500	
	AsnLeuMetValSerValThrGluProGlyGlyThrGlyThrAlaGlyAlaValAspGlyPheAspValGlyAlaLysThrGlyThrAlaArgLysPheValAsnGlyArgTyrAlaAsp		
A)	AATCTGATGGTTTCCGTAAACGAGCGGGCGGCACCGGTACGGCGGGTGGCGGTGAGCGGTTTCGATGTCGGCGCTAAACACGGGCAGGGCGCGCAAGTTCGTCAACGGCGGTTATGCGCAC	1560	
B)	-----		
C)	-----	Leu Val	
D)	-----	AC-G T-C-C-TG	
E)	-----	Leu Val	
	520	540	
	AsnLysHisValAlaThrPheIleGlyPheAlaProAlaLysAsnProArgValIleValAlaValThrIleAspGluProThrAlaHisGlyTyrTyrGlyGlyValValAlaGlyPro		
A)	AACAAACACGTGCTACCTTTATCGGTTTGGCCCCGCGCAAAACCCCGTGTGATTGTGGCGGTAAACCATCGACGAACCGACTGCCACGGCTATTACGGCGCGGTAGTGGCAGGGCCG	1680	
B)	-----	Asn	
C)	-----	Asn	
D)	-----	AA-TT	
E)	-----	AA-TT	
	560	581	
	ProPheLysLysIleMetGlyGlySerLeuAsnIleLeuGlyIleSerProThrLysProLeuThr AlaAlaAlaValLysThrProSer***		
A)	CCCTTCAAAAAATATGGGCGGCAGCTGAACATCTTGGGCTATCCCGACCAAGCCACTGACC...GCCGAGCGCTCAAAACACCGCTTAATCCGAGTATCAACGAGATTGTTTT	1800	
B)	-----		
C)	Val GlnVal Val Val		
D)	GT C-G-C -G- -A-TT -AAT-TT		
E)	-----		
	1870		
A)	ATGTTTCAGCAAGTTAAGCCCTTGGCTGAAACCGGCATCCCGACCTGTGCTGTGCAACCGCGGAGGGC		
B)	-----		
C)	-----		
D)	-----		
E)	-----		

Fig. 1 Nucleotide sequence of the *penA* gene, and the derived amino-acid sequence of PBP 2, of penicillin-sensitive and penicillin-resistant clinical isolates of *N. gonorrhoeae*. The sequence shown in full is that of the penicillin-sensitive strain LM306. Where only a single line of sequence is shown the nucleotide sequences of the *penA* gene of all five strains were identical. In other regions the sequence of the *penA* gene of LM306 is shown in line (a) and the alterations in the sequences in FA19, CDC84-060418, CDC77-124615 and CDC84-060384 are shown respectively in lines b-e.

Methods. Chromosomal DNA of *N. gonorrhoeae* strain CDC77-124615 was partially digested with *Clal* and *SauIIIa* and 5-15 kilobase (kb) fragments were isolated and inserted into pBGS131 (ref. 17) cut with *AccI* and *BglII*. Transformants of *E. coli* TG1 containing recombinant plasmids were pooled in batches of 35, and miniprep DNA¹⁸ isolated from 30 pools was tested for its ability to transform the penicillin-sensitive gonococcal strain FA19 to an increased level of penicillin resistance. DNA from three of the pools could transform FA19 to increased penicillin resistance (sixfold). Plasmid DNA from each of the transformants from the three positive pools were tested for their ability to transform penicillin resistance and in each case a single plasmid was obtained that could transform FA19 to penicillin resistance at high frequency. Each of these plasmids (pB132, 133 and 134) contained overlapping inserts of gonococcal DNA that extended for different lengths (4.5-11.1 kb) from the same *Clal* site. The *penA* gene of CDC77-124615 was located to a 2.9-kb *Clal*-*HincII* fragment by testing the ability of purified fragments of the 4.5-kb chromosomal insert of pB134 to transform FA19 to penicillin resistance. The *Clal*-*HincII* fragment was sequenced by the dideoxy method¹⁹ and was found to contain a single long open reading frame which was assigned as the *penA* coding region. The translation start was identified by using the algorithm of Stormo *et al.*²⁰ and was highly significant using both the W71 and W101 matrices (scores of 160 and 31 respectively). The *penA* genes of other gonococcal strains were obtained from gene libraries constructed by cloning chromosomal DNA, digested with *Clal* and *SmaI*, between the *AccI* and *SmaI* sites of pBGS19 (ref. 17). Recombinant plasmids containing the *penA* gene were identified by DNA hybridization using the oligo-labelled²¹ 1-kb *PstI* fragment derived from within the coding region of the *penA* gene (nucleotides 169-1,226). In each case the *penA* gene was isolated on a 4.6-kb *Clal*-*SmaI* fragment which was subcloned into both M13mp18 and M13mp19 phage for dideoxy sequencing. The sequences of the *penA* genes were determined on both strands using two sets of oligonucleotides that primed dideoxy-sequencing reactions at intervals along each strand of the gene.

PBP is predicted to extend from approximately residues 260–581, and Ser 310 can be identified as the active-site residue that is acylated by β -lactam antibiotics¹³.

The coding region of the *penA* gene of two penicillin-sensitive isolates (FA19, isolated in the USA in the 1970s; MIC of $0.008 \mu\text{g ml}^{-1}$, and LM306, isolated in 1987 in Brighton; MIC of $0.004 \mu\text{g ml}^{-1}$) differed by only three nucleotides and resulted in one alteration in the sequence of PBP 2 (His versus Asn at residue 541).

The *penA* gene of the CMRNG isolate CDC77-124615 (isolated in the USA in 1977; MIC of $2 \mu\text{g ml}^{-1}$) differed from that of the penicillin-sensitive strains at 16 positions resulting in five alterations of the amino-acid sequence. Unexpectedly, an extra codon was found in the *penA* gene of CDC77-124615 (and in other CMRNG strains; Fig. 1). The *penA* gene of a second penicillin-resistant strain, CDC84-060384 (MIC of $4 \mu\text{g ml}^{-1}$), which was isolated in 1983 during an outbreak of CMRNG infections in North Carolina¹⁴, was identical in sequence to the *penA* gene of CDC77-124615, except for an alteration at nucleotides 1678/1679 which resulted in the substitution of Pro 551 (CCG) in PBP 2 of the penicillin-sensitive strains, by Ser (TCG) in CDC77-124615, and by Leu (CTG) in CDC84-060384. Although the CMRNG strains CDC77-124615 and CDC84-060384 were isolated several years apart, the sequences of their *penA* genes show that they are derived from the same penicillin-resistant clone.

The strain CDC84-060418 (MIC $1 \mu\text{g ml}^{-1}$), which was isolated in Boston in 1984, is different from the majority of CMRNG strains in that it possesses a PBP 2 with extremely low affinity for penicillin, but has relatively minor alterations to the affinity of PBP 1 (ref. 3). The sequence of the *penA* gene of CDC84-060418 was identical to that of the other two CMRNG strains from nucleotides 1–1592, but diverged substantially throughout the remainder of the gene, and in the region immediately downstream of *penA*. The amino-acid sequence of PBP 2 of this strain differed from that of the penicillin-sensitive strain LM306 at ten positions (two insertions and eight substitutions).

CDC84-060418 also appears to be related to CDC77-124615 and CDC84-060384 as the codon insertion, and all of the substitutions (with the exception of the codon 551 change), that distinguish the *penA* gene of these latter strains from the two penicillin-sensitive strains are also found in the *penA* gene of CDC84-060418. Strain CDC84-060418, however, has numerous additional nucleotide substitutions, and a further codon insert, compared to the other resistant strains. A more detailed study of the molecular epidemiology of CMRNG strains using oligonucleotide probes will be reported elsewhere (C. G. Dowson *et al.*, manuscript in preparation).

As expected from laboratory studies, the development of altered forms of PBP 2 with decreased affinity for penicillin has occurred by multiple amino-acid substitutions (and unexpectedly by amino-acid insertions) within the penicillin-sensitive transpeptidase domain^{15,16}. The contribution of each amino-acid substitution (and insertion) to the decreased affinity of PBP 2 of the CMRNG strains for benzylpenicillin is at present unknown.

The distribution of nucleotide substitutions between the *penA* genes of CMRNG and penicillin-sensitive strains was unexpected. This was particularly striking in strain CDC84-060418 where the sequence was almost identical to that of the penicillin-sensitive strains throughout the region encoding the putative transglycosylase domain, and the amino-terminal part of the transpeptidase domain (nucleotides 1–1,592), but diverged considerably within, and downstream of, the coding region for the carboxy-terminal part of the transpeptidase domain (nucleotides 1,593–1,870). The high degree of localized sequence divergence is not directly related to the altered affinity of the PBP for penicillin as the majority of the nucleotide changes do not alter the amino-acid sequence. Its presence, however, raises the possibility that the development of low-affinity forms of PBP 2 in

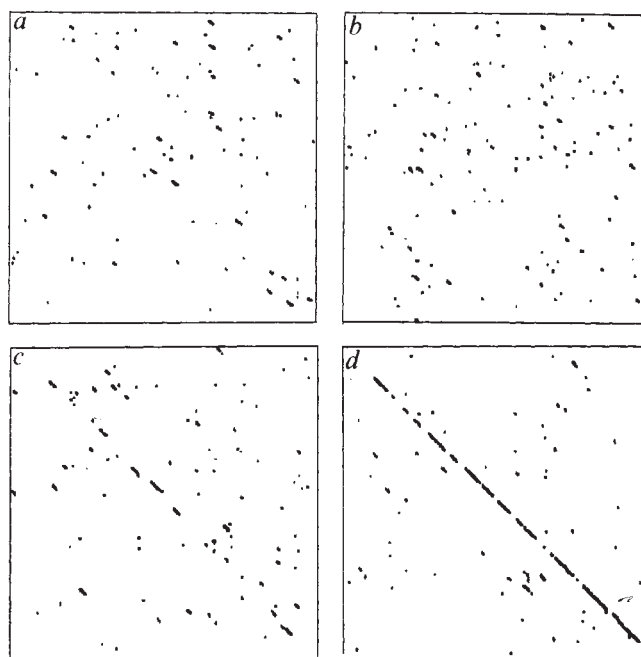


Fig. 2 Similarity of PBP 2 of *N. gonorrhoeae* to the high molecular mass PBPs of *E. coli*. Each panel shows a similarity matrix obtained using the program DIAGON²², with PBP 2 of *N. gonorrhoeae* along each horizontal axis, and PBP 1A (a), PBP 1B (b), PBP 2 (c) and PBP 3 (d) respectively along the vertical axes. A span length of 11 amino acids was used and segments having a score >132 were plotted.

CMRNG strains has involved, in addition to amino-acid substitutions and insertions, the replacement of the region encoding part of the transpeptidase domain with the corresponding region from the homologous PBP of a closely related bacterial species. This mechanism could provide the gonococcus with a more penicillin-resistant form of PBP 2 if the coding region for the transpeptidase domain was obtained from a closely related species that produced a form of PBP 2 which, by virtue of slight variations in its amino-acid sequence, possessed a lower affinity for penicillin than its gonococcal homologue.

The development of PBPs with substantially decreased affinity for β -lactams in the laboratory requires the sequential introduction of multiple amino-acid substitutions that allow the enzyme to re-model the active centre of its transpeptidase domain so that it can discriminate between the binding of antibiotic and its structurally analogous peptide substrate¹⁶. The production of hybrid PBPs by the replacement of segments of PBP genes with the homologous regions from closely related species, particularly in bacteria that are naturally transformable, may provide a more rapid route to the development of PBPs with decreased affinity for β -lactams.

I thank the Wellcome Trust for a Research Leave Fellowship, Tom Dougherty and Gordon Nichols for gonococcal strains, Jon Saunders for advice on handling gonococci, Martin Buck and Jim Brannigan for synthesising oligonucleotides, Jenny Broome-Smith for comments on the manuscript, and Agnes Hutchison for technical assistance. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00303.

Received 10 December 1987; accepted 1 February 1988.

1. Dougherty, T. J., Koller, A. E. & Tomasz, A. *Antimicrob. Agents Chemother.* **18**, 730–737 (1980).
2. Barbour, A. G. *Antimicrob. Agents Chemother.* **19**, 316–322 (1981).
3. Dougherty, T. J. *Antimicrob. Agents Chemother.* **30**, 649–652 (1986).
4. Cannon, J. G. & Sparling, P. F. A. *Rev. Genet.* **38**, 111–133 (1984).

5. Faruki, H. & Sparling, P. F. *Antimicrob. Agents. Chemother.* **30**, 856-860 (1986).
6. Broome-Smith, J. K., Edelman, A., Yousif, S. & Spratt, B. G. *Eur. J. Biochem.* **147**, 437-446 (1985).
7. Asoh, S. et al. *Eur. J. Biochem.* **160**, 231-238 (1986).
8. Nakamura, M. et al. *Molec. gen. Genet.* **191**, 1-9 (1983).
9. Spratt, B. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2999-3003 (1975).
10. Spratt, B. G. *J. gen. Microbiol.* **129**, 1247-1260 (1983).
11. Hedge, P. J. & Spratt, B. G. *FEBS Lett.* **176**, 179-184 (1984).
12. Ishino, F. & Matsuhashi, M. *Biochim. biophys. Res. Commun.* **101**, 905-911 (1981).
13. Keck, W., Glauner, B., Schwarz, U., Broome-Smith, J. K. & Spratt, B. G. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1999-2003 (1985).
14. Faruki, H., Kohmescher, R. N., McKinney, W. P. & Sparling, P. F. *New Eng. J. Med.* **313**, 607-611 (1985).
15. Hedge, P. J. & Spratt, B. G. *Eur. J. Biochem.* **151**, 111-121 (1985).
16. Hedge, P. J. & Spratt, B. G. *Nature* **318**, 478-480 (1985).
17. Spratt, B. G., Hedge, P. J., te Heesen, S., Edelman, A. & Broome-Smith, J. K. *Gene* **41**, 337-342 (1986).
18. Holmes, D. S. & Quigley, M. *Anal. Biochem.* **114**, 193-197 (1981).
19. Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).
20. Stormo, G. D., Schneider, T. D., Gold, L. & Ehrenfeucht, A. *Nucleic Acids Res.* **10**, 2997-3011 (1982).
21. Feinberg, A. P. & Vogelstein, B. *Anal. Biochem.* **137**, 226-227 (1984).
22. Staden, R. *Nucleic Acids Res.* **10**, 2951-2961 (1982).

An ATPase with properties expected for the organelle motor of the giant amoeba, *Reticulomyxa*

Ursula Euteneuer, Michael P. Koonce, K. Kevin Pfister* & Manfred Schliwa

Department of Zoology, University of California, Berkeley, California 94720, USA

The rapid, vectorial, microtubule-associated transport of organelles is believed to be mediated by specific mechanochemical transducers^{1,2}. Recent studies of various metazoan cells have allowed the identification of novel microtubule-dependent translocator molecules capable of promoting microtubule gliding across glass surfaces and translocation of inert beads along microtubules³⁻¹¹. These translocators could be involved in force generation for directional organelle movements *in vivo*. Here we report the identification of a microtubule-binding protein with characteristics expected for an organelle translocator in the giant freshwater amoeba *Reticulomyxa*. This factor has an apparent relative molecular mass (M_r) of 440,000 (440K) and sediments at 20-22S in sucrose-density gradients. It binds to microtubules under conditions of ATP depletion, possesses an ATPase activity and is sensitive to ultraviolet-induced, vanadate-dependent cleavage. Although its pharmacological properties differ from those of axonemal dynein, it can be considered to be a variant of cytoplasmic dynein. The *Reticulomyxa* high-molecular-weight protein (HMWP) promotes rapid, bidirectional movement of latex beads along *Reticulomyxa* microtubules *in vitro* at an average speed of $3.6 \mu\text{m s}^{-1}$. This protein, therefore, is a likely candidate for a microtubule-dependent motor.

Reticulomyxa possesses an extensive peripheral network of reticulopodial strands containing microtubule bundles that support bidirectional organelle movements at rates of up to $25 \mu\text{m s}^{-1}$ both *in vivo*¹² and *in vitro*¹³. Latex beads incubated in a high-speed cytoplasmic extract of *Reticulomyxa* will move bidirectionally along the microtubules of a detergent-stripped network¹⁴, suggesting that *Reticulomyxa* contains a cytoplasmic translocator. To isolate this factor, we mixed extracts with purified, taxol-stabilized bovine brain microtubules in a buffer containing hexokinase and glucose to deplete the pool of ATP, and sedimented the mixture in the centrifuge. Under these conditions, a high-molecular-weight polypeptide (HMWP)

Table 1 Release of the HMWP polypeptide from microtubules under various experimental conditions

Release conditions	HMWP released (%)
5 mM Mg-ATP	76
0.5 M NaCl	62
5 mM Mg-ATP+0.5 M NaCl	80
5 mM ATP, no Mg	39
Buffer alone (no ATP)	9
5 mM AMPPNP+3 mM Mg	3
10 mM pyrophosphate	8

The material for these release experiments was prepared as described in Fig. 1 legend, with the modifications listed above. All the material of the final supernatant and the microtubule pellet was loaded onto 7.5% SDS-polyacrylamide gels, and gel lanes were scanned in a densitometer.

selectively and specifically binds to the microtubules (Fig. 1a, lane 2). SDS-PAGE shows the relative molecular mass of this polypeptide is between that of neuronal microtubule-associated proteins MAP1 and MAP2 and above that of *Chlamydomonas* dynein heavy chains (not shown). The HMWP is effectively released from the microtubules with Mg-ATP (Fig. 1a, lane 4). Typically, the HMWP constitutes 80-90% of the non-tubulin polypeptides released from the microtubules with ATP, as determined by gel densitometry. We did not detect ATP-dependent release from microtubules of any other polypeptide, including polypeptides in the range 110K-135K, as would be expected if it were kinesin³⁻⁵. Binding of the HMWP to microtubules is optimal only in the absence of ATP. ATP prevents binding even in the presence of a 20-fold excess over ATP of the unhydrolysable ATP analogue AMPPNP or of tripolyphosphate. Conversely, AMPPNP or tripolyphosphate do not affect binding in ATP-depleted extracts. Thus, the absence of ATP is both necessary and sufficient to promote microtubule binding. These findings are in contrast to studies on kinesin^{3,5,8,10,15} where AMPPNP or tripolyphosphate promotes kinesin binding to microtubules.

Release of the HMWP is most effectively achieved with Mg-ATP or high salt, whereas buffer alone (without ATP), AMPPNP or pyrophosphate are ineffective (Table 1). The *Reticulomyxa* HMWP, like dynein heavy chains, is susceptible to cleavage at a single site when irradiated with ultraviolet light (254 nm) in the presence of vanadate and ATP¹⁶, yielding two polypeptides of 240K and 200K, respectively (Fig. 1c). Intact HMWP, therefore, has an apparent M_r of ~440K.

The 'release material', of which the HMWP is by far the most prominent non-tubulin component, possesses an ATPase activity that is insensitive to azide, but is partially inhibited by vanadate at 20 μM , and strongly inhibited at 100 μM erythro-9-[3-(2-hydroxy-nonyl)]adenine (EHNA) is inhibitory only if present in a large excess over ATP (Table 2). The effects of vanadate and EHNA on the ATPase activity are similar to their effects on reactivated organelle motility in *Reticulomyxa*¹³. The addition of taxol-stabilized microtubules slightly, but measurably, increases the ATPase activity, whereas microtubules alone do not possess any ATPase activity above background.

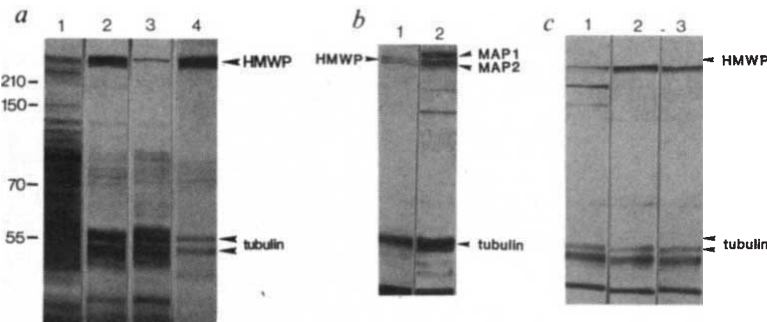
When the material released from microtubules by ATP (Fig. 1, lane 4) is sedimented in 5-30% sucrose gradients and the resulting fractions are analysed for polypeptide composition by gel electrophoresis, the HMWP sediments at 20-22S (Fig. 2a). ATPase assays of these fractions show the bulk of the ATPase activity (80-90%) to cosediment with the HMWP (Fig. 2b). We calculate the specific activity to be $0.3-0.8 \mu\text{mol min}^{-1} \text{mg}^{-1}$ protein (six experiments).

Cytoplasmic extracts, the material released from microtubules with ATP, and sucrose-gradient fractions of the release material all promote bidirectional movement of 0.2- μm diameter latex beads along stripped microtubules of *Reticulomyxa* networks.

* Present address: Department of Cell Biology, University of Texas Health Sciences Center, Dallas, Texas 75235, USA.

Fig. 1 a, ATP-dependent binding to, and release from, microtubules of the HMWP in *Reticulomyxa* cytoplasmic extracts. Lane 1, Cytoplasmic extract; lane 2, Polypeptides that bind to phosphocellulose-purified bovine brain microtubules in the absence of ATP (0.1 mg ml⁻¹ hexokinase and 10 mM glucose added); lane 3, polypeptides that bind to microtubules in the presence of 10 mM AMPPNP, no ATP added; lane 4, polypeptides released by 5 mM Mg-ATP from microtubule pellets prepared as in lane 2. The only polypeptide that binds to microtubules in the absence of ATP, and is specifically released by ATP, has an apparent M_r of >400K (arrow). Left, molecular weight markers. Gel lanes shown here are autoradiographs of a 7.5% SDS-polyacrylamide gel from an experiment with *Reticulomyxa* amoebae metabolically labelled with 35S-methionine to facilitate the detection of potential minor microtubule-binding proteins in cytoplasmic extracts. **b**, Silver-stained 6% gels of release material (lane 1) and bovine brain microtubule protein (lane 2) for comparison. The HMWP runs between MAP1 and MAP2. **c**, Ultraviolet-induced cleavage of the HMWP. Autoradiograph of a 6% gel. Lane 1, release material exposed to 254-nm ultraviolet light in the presence of 100 μ M vanadate and 100 μ M ATP for 50 min on ice; lane 2, sample treated as in lane 1, except that no vanadate was present; lane 3, untreated sample. Small arrowheads, positions of the fragments.

Methods. *Reticulomyxa* amoebae were grown in 100-mm plastic petri dishes using Rocky Mountain spring water and were fed wheat germ. For metabolic labelling, amoebae were grown for 1–2 days in 200 μ Ci 35S-methionine (Amersham) per 100-mm dish. For cytoplasmic extracts, the cell bodies were cooled for 1–2 h on ice, and rinsed 5–7 times with ice-cold spring water. The final cell pellet was homogenized in an equal volume of modified PHEM buffer²⁵ (60 mM PIPES, 25 mM HEPES, 8 mM EGTA and 0.5 mM MgCl₂, pH 7.0) with the following additives: 40 mM β -glycerophosphate, 20 μ g ml⁻¹ DNase I, 80 μ g ml⁻¹ leupeptin, 200 μ g ml⁻¹ antipain, 5 mM benzamide, 5 μ l an aprotinin solution with an activity of 0.225 U ml⁻¹ buffer, 1 mM phenylmethylsulphonyl fluoride (PMSF), 0.1 mg ml⁻¹ hexokinase and 10 mM glucose. The homogenate was spun in the cold for 25 min at 16,000g. The supernatant was removed and incubated for 20 min at 12–15 °C with an equal volume of 3 mg ml⁻¹ phosphocellulose-purified bovine brain tubulin polymerized with 10 μ g ml⁻¹ taxol, layered over a 25% sucrose cushion prepared with PHEM buffer plus 10 μ g ml⁻¹ taxol and spun at 15,000g for 60 min at 12 °C. The supernatant was carefully removed. The pellet was rinsed twice with buffer and gently loosened in half-strength PHEM buffer, pH 7.8, with 3 mM MgCl₂, 5 mM ATP and 10 μ g ml⁻¹ taxol added to release microtubule-bound proteins. Incubation was for 20 min at 12–15 °C in a shaking waterbath. The microtubules were pelleted at 15,000g for 15 min at 12 °C. The supernatant constitutes the microtubule release shown in lane 4. SDS-PAGE²⁶ was performed in either 6% or 7.5% acrylamide in the separating gel. Gels were stained with either Coomassie blue or with silver nitrate according to the procedure of Morrissey²⁷. Gels of amoeba proteins metabolically labelled with 35S-methionine were enhanced with Amplify (Amersham). Autoradiographs were obtained by exposure on Kodak XAR-5 X-ray film and developed in Kodak GBX developer. Microtubule protein was prepared from bovine brain by a modification²⁸ of the procedure of Shelanski *et al.*²⁹. Phosphocellulose column-purified tubulin was prepared essentially according to the procedure of Mandelkow *et al.*³⁰. To reduce the amount of MAPs, microtubule protein was cycled once in a buffer containing 0.5 M PIPES before applying the protein mixture to the column. The phosphocellulose-purified tubulin was >99% MAP-free as determined by quantitative densitometry of overloaded SDS polyacrylamide gels.



Of the material fractionated on sucrose gradients, only fractions that contain both the HMWP and the peak of the ATPase activity were able to support persistent bead movement at an average rate of 3.6 μ m s⁻¹ (range 2.1–7.5 μ m s⁻¹) (Fig. 2c). Movement is bidirectional in all fractions of the gradient producing bead motility. The proportion of particles moving in opposite directions varies (mean $\pm$ s.d. per fraction: 15 $\pm$ 6 anterograde, 11 $\pm$ 5 retrograde; 189 total). On average, 59% of the beads move anterograde and 41% retrograde. This movement occurs along an essentially unidirectional array of microtubules. Most (93.8 $\pm$ 6.6%) of the microtubules have their plus ends distal to the cell body (647 microtubules in three different experiments: U.E., in preparation). Other regions of the gradient fail to produce bead motility, as did coating of beads with phosphocellulose-purified tubulin, MAPs, albumin or myosin. Although most beads tend to move in either one or the other direction, we also saw examples of beads showing 'saltatory' movement, reversals in direction and acceleration or deceleration. Our attempts to use the fractions that had produced prominent bead motility in a microtubule-gliding assay are so far unsuccessful.

HMWP shows both similarities to, and differences from, other putative microtubule translocators. Like axonemal dynein, it is a high-molecular-weight, microtubule-binding ATPase whose sedimentation coefficient is similar to that of intact dynein^{17,18}. Its specific ATPase activity is in the same range as that of unactivated dynein¹⁹, and it is sensitive to ultraviolet-induced cleavage in the presence of vanadate. Whereas we observe only a slight increase in ATPase activity in the presence of intact microtubules, dynein can be activated about fivefold by microtubules¹⁹. We may not yet have optimized the conditions for microtubule stimulation, however. The HMWP differs from dynein with respect to some of its pharmacological properties, so on the basis of these preliminary data we believe that it may be a variant form of dynein.

The HMWP also differs from kinesin, a microtubule translocator from the squid giant neuron^{4,5} and other cell types^{3,10,15}. Aside from the large difference in relative molecular mass, the most distinctive characteristics are a low level of ATPase activity of unactivated kinesin (in the absence of microtubules)^{5,10,15}, and the ineffectiveness of AMPPNP to promote microtubule binding of HMWP.

Several other reports have appeared that describe proteins of M_r 295K–400K that bind to microtubules in an ATP-dependent fashion, some of which can induce microtubule or vesicle motility *in vitro*. These include a MAP2-like protein extracted from squid axoplasmic vesicles²⁰, a 400K protein present in *Caenorhabditis elegans* (nematode) egg extracts²¹, a 315K protein, termed HMW4, in spinal nerve roots²², and MAP 1C, the dynein-like translocator of bovine brain tissue²³. Although a detailed comparison between the *Reticulomyxa* HMWP and each of these proteins is not yet possible, it seems that again there are both similarities and differences. The ATPase activities of the spinal chord 315K protein and the *C. elegans* protein, for example, are markedly inhibited by low ($\leq 10 \mu$ M) concentrations of vanadate, whereas the HMWP requires much higher vanadate concentrations for significant inhibition. On the other hand, microtubule binding of the *C. elegans* protein and the *Reticulomyxa* HMWP are enhanced far more by ATP depletion than by AMPPNP. Like MAP 1C and the *C. elegans* factor, HMWP is susceptible to ultraviolet cleavage in the presence of

Table 2 Effects of various additives on the ATPase activity of the material released from microtubules with ATP

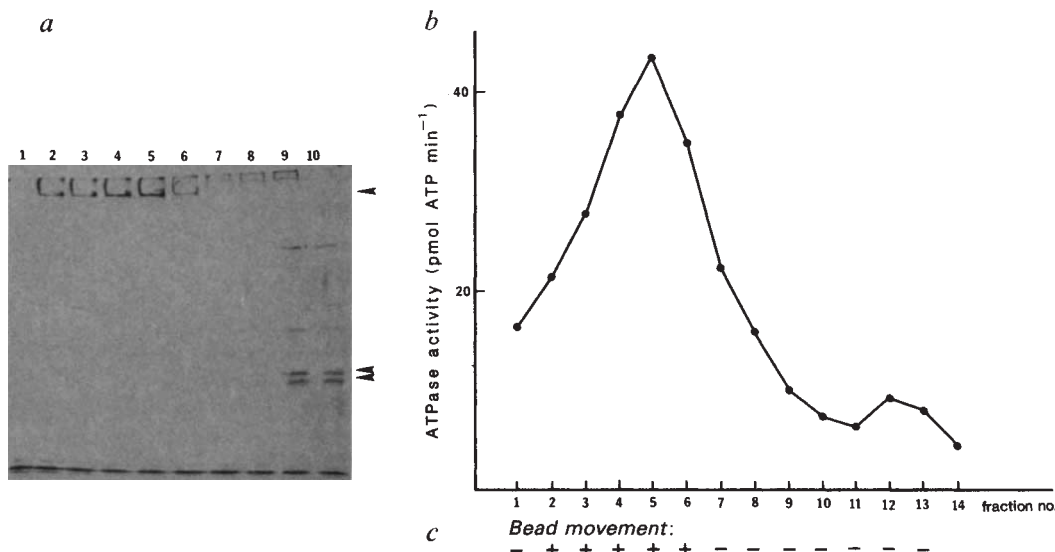
	Relative ATPase activity (%)
No additives (standard assay)	100 (4)
100 μ M vanadate	34 (3)
20 μ M vanadate	56 (2)
10 μ M sodium azide	103 (1)
2 mM EHNA + 1 mM ATP	91 (2)
2 mM EHNA + 0.1 mM ATP*	59 (1)
1 mM EHNA + 0.1 mM ATP*	61 (2)
Microtubules†	114 (3)
Microtubules alone†	— (3)

Assays were performed as described in Fig. 2 legend. Parenthesis, number of experiments.

* These assays were done with HMWP released by 0.5 M NaCl.

† Taxol-stabilized, phosphocellulose column-purified bovine brain microtubules at 1.2 mg ml⁻¹.

Fig. 2 Fractionation of the HMWP (a), ATPase activity (b), and bead-moving capacity (c) in sucrose density gradients. a, Autoradiograph of the polypeptide composition of the first 10 fractions from a 14-fraction 5–30% sucrose gradient of the material extracted from microtubules with ATP (see Fig. 1, lane 4). The positions of HMWP (arrowhead) and tubulin (double arrowhead) are indicated. Fraction numbers are shown at the top of the gel. b, ATPase activity of each fraction. There is only one main peak of ATPase activity that coincides with the peak of HMWP. c, Bead-translocating activity of sucrose gradient fractions. Only the fractions that show the presence of HMWP and the ATPase activity peak induce bead movement.



Methods. The material extracted from microtubules with Mg-ATP (Fig. 1, lane 4) was loaded in the cold onto a 5–30% sucrose-density gradient made in half-strength PHEM buffer, and sedimented at 32,000 r.p.m. in a Beckman SW50.1 Ti rotor for 15 h at 3 °C. Fractions (350–400 μ l) were collected by piercing the bottom of the centrifuge tube with a syringe needle. In control gradients, the thyroglobulin (19.2S) peak appeared in fraction 6, catalase (11.3S) in fraction 8, and bovine serum albumin (3S) in fraction 12. For SDS-PAGE, aliquots of the fractions were mixed with 4 $\times$ sample buffer and loaded onto 6% Laemmli gels. ATPase activity was assayed essentially as described by Pfister *et al.*³¹. The reaction mixture (200 μ l) consisted of sample, 1 mM or 0.1 mM ATP, [γ -³²P]ATP at 500 c.p.m./nmole unlabelled ATP, and half-strength PHEM buffer. For bead-motility assays, the fractions from the sucrose gradient were concentrated five- to sixfold using Amicon Centricon 30 filters. Polystyrene beads (0.2- μ m diameter) were added and incubated on ice for 10 min to 2 h. As substrates for bead motility, lysed and stripped microtubule arrays derived from the peripheral network of *Reticulomyxa* were used^{13,14} and visualized with the aid of an Image I image processor (Interactive Video, Concord, Massachusetts). The image processing and recording steps are as described¹³.

vanadate. Interestingly, the HMWP supports bidirectional bead movement along an essentially unidirectional microtubule array, but this finding needs to be verified in different motility assays. We have found that in a lysed cell model of *Reticulomyxa*, bidirectional transport can be switched to unidirectional, retrograde transport by cyclic AMP kinase-dependent phosphorylation²⁴, and we are now exploring a possible link between the kinase-mediated directionality switch and the HMWP.

We would like to thank Dr M. Suffness (Natural Products Branch, NIH) for the gift of taxol and Dr W. Z. Cande for critical reading of the manuscript. Supported by NIH and NSF.

Received 22 December 1987; accepted 1 February 1988.

- Schliwa, M. *Cell Muscle Motil.* **5**, 1–82 (1984).
- Vale, R. D. *A. Rev. Cell Biol.* **3**, 123–156 (1987).
- Brady, S. T. *Nature* **317**, 73–75 (1985).
- Vale, R. D., Reese, T. S. & Sheetz, M. P. *Cell* **42**, 39–50 (1985).
- Scholey, J. M., Porter, M. E., Grissom, P. M. & McIntosh, J. R. *Nature* **318**, 483–486 (1985).
- Allen, R. D. *et al. J. Cell Biol.* **100**, 1736–1752 (1985).
- Paschal, B. M. & Vallee, R. B. *Nature* **330**, 181–183 (1987).
- Porter, M. E. *et al. J. Biol. Chem.* **262**, 2794–2802 (1987).
- Kachar, B., Albanesi, J. P., Fujisaki, H. & Korn, E. D. *J. Biol. Chem.* **262**, 16180–16185 (1987).
- Cohn, S. A., Ingold, A. L. & Scholey, J. M. *Nature* **328**, 160–163 (1987).
- Mitchison, T. J. *Nature* **328**, 665 (1987).
- Koonce, M. P., Euteneuer, U., McDonald, K. L., Menzel, D. & Schliwa, M. *Cell Motil. Cytoskel.* **6**, 521–533 (1986).
- Koonce, M. P. & Schliwa, M. *J. Cell Biol.* **103**, 605–612 (1986).
- Koonce, M. P., Euteneuer, U. & Schliwa, M. *J. Cell Sci. Suppl.* **5**, 79–95 (1986).
- Kuznetsov, S. A. & Gelfand, V. I. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8530–8534 (1986).
- Gibbons, I. R. *et al. J. Biol. Chem.* **262**, 2780–2786 (1987).
- Gibbons, I. R. *J. Cell Biol. (Suppl.)* **91**, 107–124 (1981).
- Johnson, K. A., Porter, M. E. & Shimizu, T. *J. Cell Biol. (Suppl.)* **99**, 132–136 (1984).
- Omoto, C. K. & Johnson, K. A. *Biochemistry* **25**, 419–427 (1986).
- Gilbert, S. P. & Sloboda, R. D. *J. Cell Biol.* **103**, 947–956 (1986).
- Lye, R. J., Porter, M. E., Scholey, J. M. & McIntosh, J. R. *Cell* **51**, 309–318 (1987).
- Hollenbeck, P. J. & Chapman, K. J. *Cell Biol.* **103**, 1539–1545 (1986).
- Paschal, B. M., Shpetner, H. S. & Vallee, R. B. *J. Cell Biol.* **105**, 1273–1282 (1987).
- Koonce, M. P. & Schliwa, M. *J. Cell Biol.* **103**, 275a (1986).
- Schliwa, M. & van Blerkom, J. *J. Cell Biol.* **90**, 222–235 (1981).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Morrissey, J. H. *Analyt. Biochem.* **117**, 307–310 (1981).
- Euteneuer, U. & McIntosh, J. R. *J. Cell Biol.* **87**, 509–515 (1981).
- Shelanski, M. L., Gaskin, F. & Cantor, C. R. *Proc. natn. Acad. Sci. U.S.A.* **70**, 765–768 (1973).
- Mandelkow, E.-M., Herrmann, M. & Ruhl, U. *J. molec. Biol.* **185**, 311–327 (1985).
- Pfister, K. K., Fay, R. B. & Witman, G. B. *Cell Motil.* **2**, 525–547 (1982).

GUIDE TO AUTHORS

Please follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers for whom English is a second language and those who work in other fields. Please write clearly and simply, avoiding unnecessary technical terminology. *Nature's* staff will edit manuscripts to those ends if necessary. Contributors should check their proofs carefully. Because of the competition for space, many of the papers submitted for publication cannot be accepted. For this reason, and because brevity is a great assistance to readers, papers should be as brief as is consistent with intelligibility. Please note that one printed page of *Nature*, without diagrams or other interruptions of the text, has fewer than 1,300 words.

CATEGORIES OF PAPER

Review Articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance but must not exceed six pages of *Nature*.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles should be accompanied by a heading of 50–80 words written to advertise their contents in general terms, to which editors will pay particular attention. A heading (printed in italic type) is not an abstract and should not usually contain numbers or measurements. The study should be introduced in more detail in the first two or three paragraphs, which should also briefly summarize its results and implications.

Articles may contain a few subheadings of two or three words. The meaning of the text should not depend on the subheadings, whose function is to break up the text and to point to what follows. There should be fewer than 50 references.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should not have more than 1,000 words of text or more than four display items and should not occupy more than two pages of *Nature*. The first paragraph should describe, in not more than 150 words, the origins and chief conclusions of the study. Letters should not have subheadings or more than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned, most are unsolicited. They are normally between one and four pages of *Nature* in length.

News and Views articles are intended to inform non-specialist readers about a recently published advance. Suggestions should be made to the News and Views Editor. Illustrations are welcome. Proposals for meeting reports should be agreed in advance.

Scientific Correspondence is for the discussion of scientific matters, including contributions published in *Nature*. Priority is given to contributions of less than 500 words and five references. Figures are welcome.

Submission. All manuscripts may be submitted to either London or Washington. They should not be addressed to editors by name. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax (VAT).

GENERAL

All contributions submitted for publication in *Nature* should conform with these rules.

Manuscripts should be typed, double-spaced, with a good-quality printer, on one side of the paper only. Four copies are required, each accompanied by lettered artwork. Four copies of half-tones should be included. Reference lists, figure legends and so on should be on separate sheets, all of which should be double-spaced and numbered. Copies of relevant manuscripts in press or submitted for publication elsewhere should be included, clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their reference numbers.

Titles should say what the paper is about with the minimum of technical terminology and in fewer than 80 characters. Authors should avoid active verbs, numerical values, abbreviations and punctuation and should include one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name, figure number and, when known, manuscript reference number. Original artwork should be unlettered. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are permitted only if the parts are closely related, either experimentally or logically. Suggestions for cover illustrations, with captions, are welcome. Original artwork (and one copy of the manuscript) will be returned when a manuscript cannot be published.

Colour Artwork. A charge of £500 a page is made for 4-colour figures. Inability to meet these costs will not prevent the publication of essential colour figures if the circumstances are explained.

Figure legends must not exceed 300 words and ideally should be much shorter, and should use telegraphic form wherever possible. The figure should be described first, then, briefly, the method. Reference to a method described elsewhere is preferable to a full description. Methods should not be described in detail in the text.

References should be numbered sequentially as they appear in the text, followed by those in tables and finally those in the figure legends. Reference numbers apply only to papers published or in the press, and a different number should be given to each paper cited. All other forms of reference (including unrefereed abstracts) should be included in the text as personal communication, manuscript in preparation or preprint (with number and institution where appropriate). Text should not be included in the reference list. References should be abbreviated according to the *World List of Scientific Periodicals* (fourth edition, Butterworth, London, 1963–65). The first and last page numbers should be cited. Reference to books should clearly indicate the publisher and the date and place of publication.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided as much as possible and, when used, defined. Editors will shorten words if necessary. In case of doubt, SI units should be used.

Footnotes should not be used except for changed addresses or to identify the corresponding author if different from the first named.

Acknowledgements should be brief. Grant numbers and contribution numbers are not allowed.

The expanding repertoire

Simon Hunt and Jonathan Austyn

Advanced Immunology. By David Male, Brian Champion and Anne Cooke. Gower, London/Lippincott, Philadelphia: 1987. Pp.192. £25, \$39.95.

Immunology: A Short Course. By Eli Benjamini and Sidney Leskowitz. Alan R. Liss, New York/Wiley, Chichester: 1988. Pp.390. Pbk \$22.95, £15.75.

Immunology: A Synthesis. By Edward S. Golub. Sinauer/Blackwell Scientific: 1987. Pp.551. \$32.95, £26.50.

Immunology, Immunopathology, and Immunity, 4th Edn. By Stewart Sell. Elsevier: 1987. Pp.852. \$55, £37.50.

Basic Immunology: Immune Mechanisms in Health and Disease. By Stewart Sell. Elsevier: 1987. Pp.361. Pbk \$27.50, £19.50.

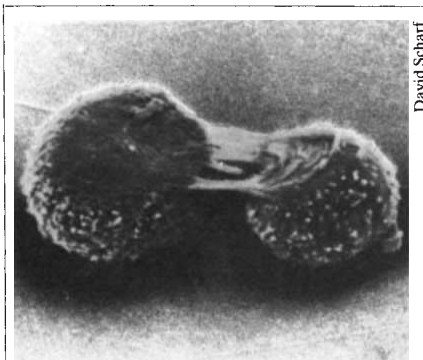
Fundamentals of Immunology, 2nd Edn. By O. G. Bier, W. Dias da Silva, D. Götze and I. Mota. Springer-Verlag: 1987. Pp.469. Pbk DM64, £23, \$29.95.

Immunology at a Glance, 4th Edn. By J. H. L. Playfair. Blackwell Scientific: 1987. Pp.79. Pbk £6.50, \$15.62.

WHAT can have stimulated a dozen immunologists to produce such a large clutch of revised and brand-new texts, all of which have appeared over the past year? To judge from the cover diagrams on two of the books, it may have been the revelation in 1984 about the molecular basis of the T-cell receptor. Many students seem happiest when the answers in their textbooks are unequivocal — nothing less than a base sequence or its derived amino-acid sequence. But, in the best of these books, the most worthy students will be nourished with provocation by yet-unsolved questions, and with a sense of how solutions may come in the future.

For the initiated, Male, Champion and Cooke make the most of their opportunity of starting with a blank screen. *Advanced Immunology* might form a natural follow-on to Roitt, Brostoff and Male's *Immunology*. After the first two (surprisingly didactic) chapters, there are three on molecules of the immune system (into which the authors unhesitatingly thrust lymphocyte development as well). Then follow four chapters on immune recognition, flipping back and forth easily between the molecular and the cellular, while a further four on interactions of immunologically active cells cover the uppers and downers of immune responses (among them is a chapter on networks in which idiotypy is very well treated). Finally, "Cellular Traffic" deals with lymphocytes and inflammatory cells. There is very little anywhere in the book on conventional immunopathology, tumour immunity, alloaggression or graft rejection, and the treatment in general is not as 'advanced' as Paul's *Fundamental Immunology* of 1984.

Nevertheless, beyond the first two chapters The Experiment is paramount; the explanations are always clear, with two-colour, multi-shade illustrations ably assisting; and the references are well chosen and up to date (to mid-1986). Because it is more recent, and the text is



Hybridoma cells, used for the production of monoclonal antibodies, in the process of dividing. The picture is taken from *Immunology: An Introduction*, 2nd Edn, by Ian R. Tizard, which appeared too recently to be included in this review. Publisher is Saunders.

more fluent, we prefer *Advanced Immunology* to its most obvious competitor, the second edition of *Immunology* by Hood, Weissman, Wood and Wilson (1984), though it does not have a comparable excitement of style nor self-test questions. It is more sophisticated and less cluttered by a historical approach than Clark's otherwise very clear and more comprehensive *Experimental Foundations of Modern Immunology*, the third edition of which appeared late in 1986. Even so, some parts such as those dealing with maturation in the thymus and control of immunoglobulin gene expression, and the now-mandatory use of CD nomenclature, should enforce a revision soon.

Also fresh, but this time for the newcomer, is Benjamini and Leskowitz's generally excellent *Immunology: A Short Course*, the declared purposes of which are to enable students to grasp the essentials to pass their course and to avoid a choking sensation in doing so — in short, to practise 'less is more'. These down-to-earth aims are largely achieved, though at some cost to stimulating the reader's curiosity, and with some confusion as to

whether mice or men should be the object of discussion.

Each of the 20 chapters corresponds to the reading assignment for a lecture, and they follow a conventional order of presentation (antigens, antibodies, complement, cellular basis, pathophysiological consequences). A glaring omission is the inflammatory process, and there is little detail on the T-cell receptor. The narrative is fluent enough and the multiple-choice questions at the end of each chapter are sufficiently thought-provoking to mean that the book can't be dismissed as a mere crammer. Italicized keywords in bold type make the student's highlighter redundant. Unfortunately, there are only seven photomicrographs in the book; the remaining illustrations are three-colour stylized diagrams, which are usually helpful, but occasionally not to be trusted. Some errors and oversimplifications will no doubt be dealt with in the next edition.

Golub's *Immunology: A Synthesis* is intermediate in level, and the most exuberant and unashamedly idiosyncratic of the present group of books, almost matching Klein's *Immunology* (1982) in this respect. The author's idols are at once the student reader and the late Richard Gershon (personifying conservatism and heresy, respectively?). He has broadened his original *Cellular Basis of the Immune Response*, published in 1981, to include the whole of conventionally-defined modern immunology. The book is very readable, as compelling as any adventure novel, Golub taking long excursions into topics such as the T locus, AIDS and immuno-detection methods. A detailed description of congenic mouse production without a nod in the direction of inbreds unbalances the otherwise masterly account of the twists and turns in the story of MHC restriction in T-B collaboration.

The book is attractively produced, in two-tone *Scientific American* style (forenames and laboratory affiliations are given, as well as surnames, and the diagrams are full of immunological iconography — a Brobdignagian syringe hovers over a Lilliputian mouse on almost every page). Often key experimental data are reproduced but the punch-lines sometimes falter, leaving the naive reader with no explanation for unclear data. There are also some serious errors — for instance "the hallmark of a T cell is Thy-1", with no qualification as to species, and there is confusion between MEL14 and "anti-MEL14". So, despite its readability and 'student-centredness', we cannot unreservedly recommend *Immunology: A Synthesis* as the sole text a student should rely on.

Sell's *Basic Immunology* is the *hors d'oeuvre* of his major text, *Immunology*,

Immunopathology and Immunity, hived off separately in paperback. Here a clinical orientation is much more apparent and the book includes, for instance, a strong section on inflammation. As an update of three previous editions, the traditional order and content of the material shows its age (for instance the prominent perpetuation of the 'T₀' cell nowadays sits uneasily; and the introduction to the MHC is via graft rejection, which is historically correct but pedagogically unhelpful), and many references are obsolete. There are quite a number of errors, and while many of the summary tables and two-tone figures are useful, some are downright confusing, especially those dealing with cellular interactions. Also, it is annoying that the index often refers to the glossary, which does not then refer to the body of the text. In fact, the best of the new edition of the complete work is the main course, immunopathology, and the sweet, immunity, which are somewhat encyclopaedic but straightforward, and offer full clinical fare not obtainable elsewhere in the books reviewed here.

In the second edition of their text, Bier *et al.* present their information in a form as densely packed as the droplets of a cold shower. This is a thorough and comprehensive account of immunology, making little concession to the explanatory duty of a textbook. English is apparently not the authors' mother tongue, and the style is slightly awkward. Some of the fault may lie with the publisher; more careful editing of the text was required, and the narrow, two-column presentation with rarely indented paragraphs is unattractive. A formal treatment has its place, but it is unlikely that students will prefer this book to the competition.

Finally, Playfair's time-tested manual has been revised in its accounts of mediators, receptors and networks, although some figures are more revised than others, which still need it. The author acknowledges that the figures "cannot be taken in at a glance". They automatically dominate and distract attention from the text. This is a pity because there is much that is fresh and helpful here, providing a measured and well-balanced perspective in remarkably few words. This is a good album for those revision students with a photographic memory.

And now, you armchair authors, stir yourselves to your keyboards. The three-dimensional structure of the MHC Class I antigen was published last September. Will we be punished with another seven textbooks to review in 1991? □

Simon Hunt is a Lecturer at the Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK, and a Fellow of Keble College. Jonathan Austyn is a Lecturer in the Nuffield Department of Surgery, John Radcliffe Hospital, Oxford OX3 1DU, UK, and a Fellow of Wolfson College.

Dealing with the technicalities

Graham J. Boulnois

Gene Cloning: An Introduction. By T.A. Brown. *Van Nostrand Reinhold*: 1987. Pp.234. Pbk £9.95, \$22.95.

From Genes to Clones: Introduction to Gene Technology. By Ernst-L. Winnacker. *VCH*: 1987. Pp.634. Hbk DM120, £47.95, \$68; pbk DM60, £23.95, \$34.

THE refinement and development of gene cloning technologies continues apace, with a proportional effect on biology and medicine. Teachers of these procedures thus have an increasingly difficult task — not only must they impart the basic information, but they must also be able to direct the student to up-to-date texts for background reading. The books by Brown and Winnacker make both jobs easier.

For complete novices, Brown's *Gene Cloning* will be essential reading. It assumes very little background knowledge and is easy to read. The problems and approaches are set out in a clear and concise way, without recourse to jargon, and the key procedures are described without reference to much of the underlying biochemistry, for which the beginner will be grateful.

A particularly attractive feature of the book is the pictorial (diagrammatic) representations of some of the practical procedures which give the reader a feel for the manipulations involved. I think, however, that some photographs of the crucial stages of the methods would have been helpful; for example, the line drawings of clear and turbid phage plaques will not mean much to those who have little practical experience of experimental biology. Nevertheless, the student (or for that matter, the more experienced researcher with no knowledge of genetic

engineering) who takes the time to read this excellent book will be in a good position to tackle the more advanced texts on gene cloning and analysis.

A particularly good example of such a book is that written by Winnacker. Although it declares itself to be an introduction, I believe that the novice will find the detail too daunting. Rather, it is students with some background in biochemistry who will be most appreciative of this book.

Winnacker has done a truly remarkable job in managing to produce a text that is at once up to date, comprehensive and comprehensible. This is particularly notable because the book was first produced in German, and then translated into English. The methods and approaches to gene cloning are described in detail and the illustrations are, by and large, well produced. The senior student will find this book useful because it sets the technologies in the appropriate biochemical context; indeed, the professional molecular biologist will find it a helpful source of those often forgotten points of chemistry and biochemistry which underpin the methods. The extensive reference lists at the end of each chapter are also a good source of information.

A fault of many books on gene cloning is that they tend to suggest that genetic engineering is a discipline in its own right, rather than a collection of particularly powerful methods. Brown and Winnacker have tried to set things to rights in two different ways. Brown has three chapters dedicated to the application of gene cloning in basic research and 'biotechnology', whereas Winnacker illustrates the applications of the methods throughout the text. Both approaches are only partially successful, but this minor failing can probably be rectified in the lecture theatre. □

Graham Boulnois is a Reader, and Lister Institute-Jenner Research Fellow, in the Department of Microbiology, University of Leicester, Leicester LE1 9HN, UK.



*The high jump — the Hoolock gibbon *Hylobates hoolock*, a brachiating ape, throws itself from branch to branch by rapid ricochetal arm swinging. Gibbons and siamangs are the only true brachiators. The picture is taken from *The Natural History of the Primates* by J.R. Napier and P.H. Napier, recently issued in paperback by The British Museum (Natural History)/Cambridge University Press. Price is £9.95. For review see *Nature* 316, 223 (1985).*

Expert differences

J. A. Howell

Modern Biotechnology. By S. B. Primrose. Blackwell Scientific: 1987. Pp. 176. Hbk £25, \$47.50; pbk £12.95, \$24.50.

Introduction to Biotechnology. By C. M. Brown, I. Campbell and F. F. Priest. Blackwell Scientific: 1987. Pp. 169. Pbk £9.95, \$22.50.

Biotechnology: The Biological Principles. By M. D. Trevan, S. Boffey, K. H. Goulding and P. Stanbury. Open University Press, Milton Keynes/Taylor & Francis, Philadelphia: 1987. Pp. 256. Hbk £32.50, \$45; pbk £14.95, \$22.

THE academic world is awash with biotechnology courses. They range from fully structured degree programmes to sets of lectures, given by a hastily assembled clutch of individuals, who are asked to explain their interests without any overall coordination of the topics. Neither sort of course should be offered to undergraduates. Although there is a case for a multidisciplinary course to be taught to postgraduates, until there is a solid background of textbooks available to demonstrate the heuristic method of the subject an undergraduate degree is premature.

All disciplines develop their own *modus operandi* and unique problem-solving techniques. The techniques may be logical, mathematical, empirical, knowledge-seeking, analytical or design-orientated, but in each case the technique fits the field. Microbiologists are acutely aware of the diversity of life; every organism is different and it is extremely difficult to draw generalizations. With biochemistry, some general principles are apparent, but prediction in a new environment is no easy matter. In physics and engineering, general principles dominate and are reliable, thus allowing prediction of the results of experiments and extension of results to different scales and different materials. Chemistry solves problems by creating new materials or simplifying pathways; a knowledge of the past allows synthesis of the future. Design engineers behave much like chemists, creating the new from the principles of the old to solve problems that are new or have currently unsatisfactory solutions.

Where, in all this, does biotechnology lie? It is not yet a discipline but a field of application to which many disciplines contribute. To the biochemist, biotechnology is principally concerned with developments in recombinant DNA technology, as this is the source of new products and requires many new techniques. The older technologies such as brewing tend to be dismissed. On the other hand, to produce new materials industrially through biochemical engineering requires old

technologies, use of established methods in the discipline and a small but vital contribution of novel technology. Biotechnology is not an isolated endeavour, but is grafted onto the old technology in which its practitioners must be proficient.

These three textbooks cover a wide range of topics, but only one (Primrose's) might serve as a text in a future biotechnology degree course. The others might be used in connection with single-term or two-term courses taken as an option within another discipline's degree programme.

As a biochemical engineer, I found it interesting to read the three books and see that in their coverage of process engineering there is all too little about the essential predictive design aspects which are fundamental to this discipline. The microbiologist or biochemist in biotechnology will rarely become a process engineer, but he will certainly have to work with one. This multidisciplinary collaboration will never be fruitful until practitioners in each discipline understand their counterparts' approach to problem solving.

It is a pity that the engineer's approach was not communicated in these books. The engineer, fortunately, has the flexi-

bility to understand the biologist's way of doing things, so will the books be useful to engineers wishing to learn how the biologists tackle problems? Primrose's *Modern Biotechnology* certainly will. It is readable and emphasizes how the molecular geneticist uses the array of new techniques to solve the problems. On the other hand, it may be too advanced for anyone below the final year of an undergraduate course in biochemical engineering. The other two texts are pitched more modestly. That by Brown *et al.* covers a wider field, being orientated more towards areas of application, whilst that of Trevan *et al.* deals with the techniques. Both will be valuable in providing background reading and background information for a course for chemical engineers, but because they are long on facts and shorter on principles they will be less useful as formal texts. But they can be used to flesh out lectures on principles, and allow students to have one book to cover the range of biotechnology which has normally only been possible with a collection of texts. □

J. A. Howell is a Professor in the Department of Chemical Engineering, University of Bath, Claverton Down, Bath BA2 7AY, UK.

Finding a niche

Colin Townsend

Community Ecology: Pattern and Process.

Edited by Jiro Kikkawa and Derek J. Anderson. Blackwell Scientific: 1987. Pp. 444. Hbk £39.50, \$56; pbk £19.80, \$35.

MULTI-author volumes can be expected to sum up how a particular discipline stands — reviewing and synthesizing principle and practice, and providing pointers for future research. Several such books on community ecology have appeared in recent years. *Ecological Communities*, edited by Strong *et al.* and published in 1984, conveyed the sense of excitement and, indeed, conflict between proponents of different viewpoints that is evident in a dynamic and topical area. Diamond and Case's *Community Ecology*, a collection of essays that appeared in 1986, provided an excellent review of a maturing though still lively science. Most recently, Gee and Giller's *Organization of Communities* has continued the process by filling in some taxonomic and habitat gaps, and stressing insights provided by palaeoecology. The present book has some stiff opposition.

Ecologists ask whether communities are mere unstructured assemblages or tightly linked groups of interacting species. Moreover, if repeated patterns in community organization are discernible, what underlying factors are responsible? Most of the chapters in this multi-author textbook are firmly in the descriptive

phase, searching for patterns and beginning to suggest hypotheses. Several of the early contributions have an old-fashioned feel. Others are excellent undergraduate-level reviews, but contain little that is new.

In a few places, the reader is taken beyond mere description to the generation and testing of hypotheses. Terborgh and Robinson discuss convergence in community structure in different parts of the globe. Lawton and MacGarvin look for patterns and explanations about the organization of communities of insect herbivores as compared to mammalian herbivores. And Schoener evaluates tests of the role of competition in structuring communities. The largely anecdotal chapters on prey-predator interactions and succession are disappointing, as they fail to present much of the extensive experimental work on these topics. Sadly, in one of the few areas of community ecology where experiments have been attempted (competition), Underwood delivers a searching and scathing critique of the lack of attention to proper experimental design.

The good parts of this book accurately reflect the problems in community ecology — the anecdotal nature of some of the evidence, the difficulty of developing a coherent body of theory given such a complexity of issues, and the problems in designing experiments. But it generally fails to convey the excitement and vitality being generated by many good scientists working in this important area. □

Colin Townsend is a Lecturer in the School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Inherited wisdom

J.G.M. Shire

Genes III. By Benjamin Lewin. Wiley: 1987. Pp.761. Hbk £33.45, \$38.95; pbk £16.95, \$17.50.

Essential Genetics. By Peter J. Russell. Blackwell Scientific: 1987. Pp.493. Pbk £13.50, \$29.95.

Genetics. By Geoffrey Zubay. Benjamin/Cummings: 1987. Pp.1,049. \$39.95, £19.95.

GENETICS is concerned with questions about the mechanisms of inheritance: what carries inherited information, how information is passed between generations and between populations, how it is organized and how it brings about its effects. People with knowledge of genetics can contribute to many aspects of biology, from the ecological to the cellular and biochemical levels of analysis, but to do so they need to know about the breadth of the subject as well as having specialized knowledge within it. The remarkable advances in all areas of genetics pose a problem for today's students: how are they to see the wood for the burgeoning trees?

Although all of the three books reviewed here have similar titles, they set about helping the student in quite different ways. In *Genes III*, Lewin restricts himself to DNA. He describes its structure, replication and expression, together with the control mechanisms involved. The result is a detailed, even encyclopaedic, account of the currently agreed facts and hypotheses in these areas. The derivation of conclusions from actual experimental data is shown only rarely. Ordered tetrads, for example, are used to demonstrate the abnormal pattern of spores found following gene conversion and post-meiotic segregation, but without any indication of the normal spore pattern following an ordinary crossover between gene and centromere. The first 20 pages cover mutagenesis, transmission genetics and gene mapping, but are far too condensed to be useful to students.

Students reading *Essential Genetics* will have little difficulty gaining a sound understanding of inheritance in both pro- and eukaryotes. They will also be introduced to the way that DNA is organized, how its information is expressed, and to how genes behave in populations, as well as being given insight into the genetics of their own species. The balance of coverage is about right, although fungal genetics are covered in too much detail whilst the section on developmental genetics would have been better for reference to sequence reorganization and differential promotion and splicing of mRNA. Throughout the text experimental data are elegantly used to indicate how conclu-

sions are logically derived from observations. What happens during crossing-over is clearly shown in Russell's book, but readers of *Genes III* will assume that exchange of pairing partners precedes chromatid breakage and reunion. Students who have read *Essential Genetics* will be well prepared for advanced courses on specialized aspects of genetics, and they will be well able to set those aspects in context.

Zubay's *Genetics* follows a tradition, notably successful in the past, of combining breadth and depth. The resulting 1,000 pages, even with input from 15 co-authors, is a compromise. Despite the use of colour, bold-type and underlining, and the inclusion of questions and answers,

students will find it hard to read the book because of the excessive amount of detail given. Even so, only four advanced topics are covered: animal development, evolution, mobile genetic elements and cancer viruses. Keeping these up to date will require frequent revision, probably on the two-yearly cycle of *Genes*.

As Lewin recognizes in his epilogue, the recent history of genetics has been full of surprises. The book published in 1987 that will best help students prepare for such unpredictable events is *Essential Genetics*, for it shows most clearly the assumptions that underlie current explanations. □

J.G.M. Shire is a Professor in the Department of Biology, University of Essex, Colchester CO4 3SQ, UK.

Further selection

Joe Felsenstein

A Primer of Population Genetics, 2nd Edn. By Daniel L. Hartl. Sinauer/Blackwell Scientific: 1988. Pp.305. \$15.95, £12.15.
Population Genetics: Basic Principles. By Donald P. Doolittle. Springer-Verlag: 1987. Pp.264. DM70, £26, \$35.

OF similar length, title and degree of detail, and aimed at students of a similar level, Hartl's and Doolittle's books are designed for very different audiences. Hartl intends to integrate data and theory for natural populations, with emphasis on single-locus polymorphisms and molecular evolution. Doolittle aims at providing the foundation of population genetic theory for quantitative genetics in plant and animal breeding. Having these quite different niches these books are scarcely in competition, and each should be successful.

Hartl's book is a revision and updating of a well-received and popular first edition, and I think it is a considerable improvement. The present organization seems clearer than in the original, and many of the data examples on molecular polymorphisms, molecular evolution and quantitative genetics are new. The book is divided into four major sections. The section on genetic variation covers gene frequencies and polymorphisms, as well as the theory of random mating, assortative mating and inbreeding. 'Causes of Evolution' covers the evolutionary forces of random genetic drift, mutation, migration and selection, mostly as theory but with data examples.

The largest body of new material is in Chapter 3 on molecular population genetics, which introduces neutral theory and molecular clocks, but also briefly describes patterns of nucleotide substitution and the evolution of mitochondria, chloroplasts, multigene families and

transposable elements. Chapter 4 covers the genetics of quantitative characters. This material replaces some of the coverage of ecological genetics that appeared in the first edition.

While Hartl achieves a nice balance between theory and data, derivations of the details of the theory are quite limited. Students will not necessarily be able to tell where the theory actually comes from, and sometimes the theoretical results are imprecisely stated. For example, the formula for the effective breeding size of a population whose size fluctuates is an approximation, but this is never indicated, and Fisher's 'fundamental theorem of natural selection' is presented without making it clear that it is inexact in the presence of dominance effects.

Doolittle covers the algebraic theory in more detail than does Hartl, and without attempting a comparable amount of coverage of data. For my taste there are too few diagrams illustrating the theory. The book is divided into 44 short chapters called 'lectures' and into five major parts. The order of presentation of theory is similar to Hartl's, but with much more coverage of quantitative genetics. Each lecture is followed by one or a few exercises, whose solutions are given and discussed at the end of the book, which also includes an appendix with review lectures on basic genetics, probability and statistics.

While fairly detailed, Doolittle's book is not intended as an advanced textbook of quantitative genetic theory — there is no coverage of such topics as index selection or maximum likelihood estimation of variance components.

Both of these books are successful in achieving their authors' stated aims. Hartl's ought to be received even more enthusiastically than was his first edition, and Doolittle's will be a useful theoretical supplement to Falconer's classic textbook on quantitative genetics. □

Joe Felsenstein is a Professor in the Department of Genetics, University of Washington, Seattle, Washington 98195, USA.

Intrepid algae

Ralph A. Lewin

Introduction to Phycology. By G. Robin South and Alan Whittick. *Blackwell Scientific*:1987. Pp. 341. Hbk £32.50, \$39.95; pbk £15.95.

THESE days, you have to be both very brave and very knowledgeable to write a book on phycology — even an introduction to the subject. The reason is that the algae are such a diverse group, in many ways much more so than, say, land plants. Algae live deeper in the seas or grow higher up mountains, further north or south, in hotter springs, drier deserts, saltier or more acidic pools, than any land plant.

The algae include both prokaryotes and eukaryotes. Some make skeletons of incredibly delicate or grotesque tracery out of silica or lime. Some have three dissimilar generations in their life-cycles, in contrast to most land plants (which have two) and most animals like ourselves (which have only one). Some algal classes evolved even before the invention of the enneanema (the cilium or flagellum of nine-fold symmetry that characterizes all animals and almost all other plants). Some have chromosomes that never disappear, even in resting-stage nuclei. Some have only one nucleus, though their plant cell may be as long as a pencil. Some have no cells at all, their nuclei slithering to and fro in tube-like labyrinths. Some algal flagellates harpoon prey larger than themselves, haul it in by net, and quickly digest it by extracellular enzymes. Some

parasitic forms invade foreign tissues by insinuating their own nuclei into host cells. Some autotrophs photosynthesize by using plastids evidently stolen from unrelated types; in others, plastids of different origins battle for supremacy in a common cytoplasm.

It is a tough assignment to know about such wonders and then write about them in a way that will interest even blasé undergraduates. South and Whittick have succeeded pretty well in measuring up to the task. Their book is liberally illustrated with graphs and charts, useful tables and clear diagrams and photographs, and exhibits a reasonable balance of discussion of taxonomy, cytology, morphology, physiology, biochemistry, ecology, phylogeny and applied phycology.

There is an adequate index, and for those whose enthusiasm is whetted to seek more details there are 30 or so pages of fully titled references. Almost all of these are to the original literature published within the past 15 years; few are to secondary reviews. In fact, much of the text is written in the laconic style of such reviews, and I suspect that the concise syntax (and the absence of a glossary) may make the book rather hard going for the neophytes for whom it is presumably intended. But, if students plough through manfully, the effort will certainly be worthwhile.

On the whole I consider this as good an introduction to the subject of phycology as exists in the English language. Indeed, it's a strong contender for the title of The Best So Far. □

Ralph A. Lewin is a Professor of Marine Biology at the Scripps Institution of Oceanography, University of California, La Jolla, California 92093, USA.

The ways of being spineless

Alan Brafeld

Animals Without Backbones, 3rd Edn. By Ralph Buchsbaum, Mildred Buchsbaum, John Pearse and Vicki Pearse. *University of Chicago Press*:1987. Pp. 572. Hbk \$25, £19.95; pbk \$17, £11.95.

Living Invertebrates. By Vicki Pearse, John Pearse, Mildred Buchsbaum and Ralph Buchsbaum. *Boxwood Press*:1987. Pp. 848. Distributed by *Blackwell Scientific*, \$49.95, £29.

Invertebrate Zoology, 5th Edn. By Robert D. Barnes. *Holt, Rinehart and Winston*:1987. Pp. 893. \$33, £33.

INVERTEBRATES fascinate us largely because they are nothing like us. There are about 30 phyla of invertebrates, each built on a body plan quite distinct from the others and from the vertebrates. Some have no gut, some no heart (or even blood), some no brain to speak of, some no legs. Yet, in their very different ways, all of them are effective and viable.

Studying such varied solutions to the challenges that face all animals is enriching and rewarding, and so every undergraduate biologist learns something of invertebrate zoology. In recent years there has been a shift of emphasis in the teaching of the subject, away from highly detailed study of comparative anatomy and systematics and towards more unifying and exciting aspects such as physiology, ecology and behaviour. This is a response to increased knowledge and to a tendency towards synthesis rather than analysis. It is more fun, now that we have such a lot of information, to find gratifying similarities rather than detailed differences. Solid tomes on invertebrates are still necessary, but the better ones will reflect this trend.

Animals Without Backbones first appeared in 1938 and has been going strong ever since. Four members of the Buchsbaum family have got together to write the latest version, which is a mosaic of old and new. There has been extensive updating in all directions (the last revision was in 1975) but a lot of the drawings and photographs will be familiar to the many owners of the classic 1948 edition. Perhaps too many of the old illustrations have been retained (there is still that dreadful drawing of the nerve net of a hydra) but a lot of greatly improved artwork makes the general quality much higher than before. There are also hundreds of well-printed photographs interspersed among the text and these give the reader a good feel for what the animals are really like, something which even the best drawings would not have been able to achieve. ▶



Get weaving — a ribbon worm of the phylum Nemertea. The picture is taken from a colour original in *Biology: The Unity and Diversity of Life*, 4th Edn, by Cecie Starr and Ralph Taggart, just published by Wadsworth, California. Other recent editions of basic biology textbooks received in the Nature office include *Life: The Science of Biology*, 2nd Edn, by William K. Purves and Gordon H. Orians (Sinauer); *Biology*, 3rd Edn, by Karen Arms and Pamela S. Camp (Saunders); *Biology*, 2nd Edn, by Leland G. Johnson (William C. Brown, Iowa); and *Biosphere: The Realm of Life*, 2nd Edn, by Robert A. Wallace, Jack L. King and Gerald P. Sanders (Scott, Foresman & Co., Glenview, Illinois).

The minor phyla still get scant attention and the chapter on phylogeny is too brief to say anything much of worth, but that on fossils is greatly improved by the many photographs. The chapter which considers how and where zoologists carry out their research and communicate their findings will give the novice an idea of how the information given elsewhere in the book came to light. There is a short section on classification, and a list of books and journals.

The same four authors have produced *Living Invertebrates*, a more comprehensive work designed for a more advanced readership. Many of the drawings by Mildred Waltrip appear in both books, and helpful photographs abound (with 35 colour pages illustrating the chapter on pigments and colouration). A book list is included, but again there are no references to original papers in journals.

Unfortunately, the chatty and rather naive style of writing, which works well at the level of *Animals Without Backbones*, is here too. In the context of this larger and more serious-minded book one can easily find it irritating. At the level at which this book is pitched, it is not a virtue to avoid technical terms wherever possible — such words have a precise meaning and therefore a very real value. Students are quite capable of mastering a remarkable number of them, particularly if they learn their etymology, and it is impoverishing the language to shun them. So the book falls between two stools, with neither the freshness and simplicity which makes *Animals Without Backbones* so appealing to the newcomer to invertebrates nor the depth of treatment and scholarship shown by Barnes's *Invertebrate Zoology*.

Now in its fifth edition, *Invertebrate Zoology* has always been a very good book for the serious student, for it contains a great deal of information, is well illustrated and has plenty of references. Each edition was getting ever more weighty, however, and I gently took the author to task in these columns when considering the fourth edition. Apparently it was generally felt to be too large and a reduction of about 170 pages has been made. New material is included, however, and there are two distinct improvements: an expanded glossary combined with the index, and 'boxed essays'. These are short treatments of topics which recur in various invertebrate groups (such as trochophore larvae, eyes, nephridia, bioluminescence) and they provide useful overviews and comparisons.

So Barnes is better than ever, though it's still the case that only 3 per cent of the space is given to the insects. But then, nobody's perfect. □

Alan Brafield is a Senior Lecturer in the Department of Biology, King's College London (KQC), Campden Hill Road, London W8 7AH, UK.

Organic evolution

John Mann

Organic Chemistry. By K. Peter C. Vollhardt. W.H. Freeman: 1987. Pp. 1,275 plus answers to exercises and index. Hbk \$49.95, £49.95; pbk \$33.95, £18.95. (*Study guide and solutions manual*, pbk \$20.95, £16.95).

Organic Chemistry. By Francis A. Carey. McGraw-Hill: 1987. Pp. 1,219. \$45.95, £43.95.

Basic Organic Chemistry: A Mechanistic Approach, 2nd Edn. By J. M. Tedder and A. Nechvatal. Wiley: 1987. Pp. 303. Hbk £33, \$69; pbk £9.95, \$19.95.

The Art of Problem Solving in Organic Chemistry. By Miguel E. Alonso. Wiley: 1987. Pp. 324. £25.95, \$29.95.

THE standard textbook of organic chemistry has been evolving continuously for the past 30 years. Like the movies, it began in black and white, and has progressed stage by stage until now it is multicoloured. In the 1960s we had the arrival or new editions of the excellent books by the Fiesers (*Advanced Organic Chemistry*) and Finar (*Organic Chemistry*), and the somewhat dull and factual *Organic Chemistry* by Morrison and Boyd.

The latter two books, currently in their sixth editions, are still with us, and although Morrison and Boyd now includes colour, Finar has retained its classical monochrome appearance. One feature that raised the Fiesers' book (and to a lesser extent Finar's) above all others, then and since, was that it read like a novel. All the essential facts were reported, but interspersed with short biographies of famous chemists, and interesting pieces on the industrial relevance of the topics discussed. The book was brim full with the excitement of organic chemistry.

In the intervening years much of this excitement has gone from textbooks, to be replaced by lists of reactions and masses of problems; and all the 'advanced' topics are confined to boxes or appendices where they cannot disturb the learning process. The market for these books is nonetheless enormous, and there is a seemingly annual competition to produce the new text with the most novel features.

In 1987, Peter C. Vollhardt's *Organic Chemistry* wins hands down. It employs four colours to simplify reaction mechanisms, to identify signals in spectra and to clarify functional-group interconversions; and the depictions of three-dimensional structures (and of orbitals) are not only clear but very appealing. The content is much the same as that of other recent textbooks, though there is a strong bias towards organic synthesis, and the applications of organic chemistry to the biological sciences and to industry are

stressed. One nice touch is that all of the reactions reported were checked in the literature, or were run in the laboratory to confirm their authenticity. There is also a good chapter on heterocyclic chemistry.

Overall, the book is visually appealing, up-to-date and, although it does not read like a novel, it is informative and easy to follow. It is hard to imagine next year's novelty, though three-dimensional 'pop-up' structures might be one possibility!

It is rather unfortunate that Francis A. Carey's book has to be reviewed at the same time as Vollhardt. It, too, is eminently suitable for undergraduate classes, has plenty of organic synthesis and other modern features; but only three colours are employed, and they are not used as imaginatively as in Vollhardt's book. Carey has some advantages over Vollhardt, in that classes of compounds (such as organometallic reagents and isoprenoids) are dealt with all together, and are not dispersed through several chapters. But, overall, the consumer is likely to choose Vollhardt for its sheer good looks.

British textbooks of organic chemistry are a comparative rarity, and so the second edition of Tedder and Nechvatal's *Basic Organic Chemistry* is to be welcomed. The first edition appeared 20 years ago, and one of the main developments since then has been the increasing application of spectroscopic techniques in organic chemistry. Most chapters of this new edition begin with a summary of the main spectral features (principally NMR and infrared) of the functional group under discussion, but otherwise the authors concentrate on basic mechanisms.

The text is surprisingly comprehensive for its 300 pages, and the authors even manage to include brief accounts of natural products, carbohydrates, polymers and molecular orbital theory. But this book will probably be mainly of use to first-year chemistry undergraduates or to biologists requiring a grounding in organic chemistry. One annoying feature is the poor quality of the printing, and in visual appeal the volume compares very unfavourably with the two previous books.

Miguel E. Alonso's *The Art of Problem Solving in Organic Chemistry* is very different from the other books reviewed here. It contains 56 problems in synthesis, with worked answers, but these are probably too difficult for most undergraduates. They will, however, be ideal for the post-graduate. The illustrations range in quality from good to confusing, but the answers are usually very helpful. This book will be an essential purchase for all PhD supervisors, allowing them to set problems without having to comb the research literature — and thus freeing them to concentrate on raising money. □

John Mann is a Lecturer in the Department of Chemistry, University of Reading, Whiteknights, Reading RG6 2AD, UK.

Various inorganic attractions

J. A. McCleverty

Basic Inorganic Chemistry, 2nd Edn. By F. Albert Cotton, Geoffrey Wilkinson and Paul L. Gaus. Wiley: 1987. Pp.708. Hbk £34.25, \$39.90; student edn £12.95, \$9.60.

Mechanisms of Inorganic Reactions. By Dimitris Katakis and Gilbert Gordon. Wiley: 1987. Pp.384. £34.50, \$39.95.

Structural Methods in Inorganic Chemistry. By E.A.V. Ebsworth, David W.H. Rankin and Stephen Cradock. Blackwell Scientific: 1987. Pp.456. Hbk £29.50, \$44.50; pbk £13.50.

THESE three books have two things in common: they are all concerned with inorganic chemistry, and they all provide problems and study guides, an invaluable aid to instructors struggling to find new topics and questions for tutorials and examinations.

Basic Inorganic Chemistry is an update of an earlier text directed at those who seek a general introduction to most, but not all, aspects of the area. It is not intended to be a 'baby' version of the *Advanced Inorganic Chemistry* by Cotton and Wilkinson, the fifth edition of which will soon appear, but a complete text, one-third of which is dedicated to general preliminaries such as structure, bonding, solids, coordination chemistry, acids and bases, and periodicity. The middle section describes the chemistry of the 'main group' elements, by individual element or group. Part three presents transition metal, lanthanide and actinide chemistry in a style similar to the *Advanced* series, by horizontal divisions. The final part of the book gives outlines of special topics — π -acceptor ligands and organometallic chemistry, catalytic uses of organometallics, and bioinorganic systems. Topics in solid state chemistry are barely touched on, which is a pity since this is a currently fashionable area. It is arguable, however, whether a single text in inorganic chemistry, to be manageable and useful, could cover everything of interest and possible relevance.

That minor criticism apart, this is a well-rounded book, providing a description of most of inorganic chemistry on a predominantly factual basis relating to the periodicity of the elements, but with deference to the presentation of this vast area as topics. This makes it somewhat different to and, in a sense, more historically rigorous than Huheey's book *Inorganic Chemistry*, for which Cotton, Wilkinson and Gaus will provide serious competition.

Katakis's and Gordon's treatment of

mechanisms of inorganic reactions is not intended for specialists but more for the informed beginner, at a level requiring knowledge of basic inorganic chemistry, thermodynamics and kinetics. It reviews the relationship between mechanisms, kinetics and equilibria and touches on the important factors influencing them. There are chapters dealing with the concept of the activated complex, experimental methods, the relationship between mechanism and structure, group-, atom- and electron-transfer processes, homogeneous catalysis and inorganic photochemistry. Each section has a sensibly limited number of examples illustrating the basic principles. It is a very useful supplementary text for a course of general inorganic chemistry which illustrates mechanistic study as being a constituent part of the subject area, rather than as an activity on its own.

The book by Ebsworth, Rankin and Cradock is a joy to read and, for practitioners as well as students and instructors, is extremely valuable. It deals with the application of nuclear magnetic, quad-

rupole and electron-spin resonance, rotational, vibrational, electronic and photoelectron spectroscopy to real inorganic problems. It also discusses the uses, in a structural context, of Mössbauer spectroscopy, diffraction techniques and mass spectrometry.

The jewel in this particular crown is the chapter on case histories, where real problems are confronted and the contents of the structure-solving tool box deployed to the greatest effect. I must comment that there could have been some additions, for example, on aspects of optical activity, magnetism and electrochemistry, since this is a book concerned with analysis in its widest sense. These areas have as much to do with structure as, say, mass spectrometry, and are all part of the serious inorganic researcher's structure-solving kit. This is not a criticism — merely an expression of greed for more of this excellent reprint. □

J. A. McCleverty is a Professor of Inorganic Chemistry in the Department of Chemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

Technical merit

T. S. West

Instrumental Analysis in the Biological Sciences. By M. H. Gordon and R. Macrae. Blackie: 1987. Pp.244. Pbk £17.95.

Introduction to Instrumental Analysis. By Robert D. Braun. McGraw-Hill: 1987. Pp.1,004. \$52.95, £45.95; student edn £14.95.

IN RECENT years chemical analysis has undergone a revolutionary change which, although familiar to the chemists and physical scientists who have been responsible for it, is less well known to biological scientists. Biologists, however, can gain enormously from these new, and generally highly sophisticated, instrumental techniques, many of which employ microprocessors to simplify and improve their operation and make them more user-friendly.

Gordon and Macrae's informative and well-balanced book succeeds admirably in providing an easy-to-assimilate account of the standard separation-determination techniques, such as liquid-phase chromatography, gas-liquid chromatography and electrophoresis, that have become indispensable in the biological laboratory. Spectroscopic techniques, which are the most widely used of all physico-chemical investigative procedures, are considered in categories such as ultraviolet and visible spectrometry, fluorescence and phosphorescence, and infrared absorption spectrometry. Nuclear magnetic resonance, electron spin resonance and elec-

tron paramagnetic resonance spectrometries are dealt with well, but in less detail than is probably warranted by their ever-increasing importance in biology. The treatment of flame emission, atomic absorption and atomic fluorescence spectrometries is thorough enough for the intended readership, even perhaps too much so, but the electrothermal atomization techniques deserved more attention in relation to their applicability to biological materials. The sections on organic mass spectrometry and electrochemical devices are brief but useful and largely to the point.

The authors could profitably have devoted more attention to the nature of the structural information that can be won by many of the techniques they consider. They might also have addressed the problems of correct sampling of biological tissues, without which all the sophisticated science may be pointless. But, in all, this is a most useful introduction to the theory and general instrumental background of modern physico-chemical investigative techniques.

Robert Braun has produced a much larger book, with different aims. At the least, he deserves full marks for effort in compiling such a detailed — sometimes almost laborious — review of modern instrumental techniques.

Like most comprehensive texts, this one has its good and its less good points. For me the introductory treatment of electronic circuitry, operational amplifiers, logic devices and microcomputers in instrumental 'innards' is very worthwhile, and the inclusion of photoacoustic spectrometry, refractometry, thermo-

chemistry and electron spectrometry is admirable. Otherwise, the text gives a sound account of most of the analytical techniques of spectroscopy, electrochemistry, radiochemistry, chromatography, mass spectrometry (but not, strangely, for mineral samples) and so on, and includes an all-too-brief review of automation in the laboratory. The book is well illustrated with graphs, line diagrams and tabular summaries, and there are problems for the student, together with answers associated with all the chapters except for the introduction and that on automation.

I found the book to be reliable and sound, rather than inspirational, which is perhaps as much as one should expect from a monolithic work of this nature. Certainly, there is little missing from it, though sampling — the first step in any analysis — is unfortunately ignored. □

T. S. West Macaulay Institute, Craigiebuckler, Aberdeen AB9 2QJ, UK, was formerly Professor of Analytical Chemistry at Imperial College and Director of the Macaulay Institute.

A whole mind of information

A. R. H. Gellatly

Cognitive Science: An Introduction. By N. A. Stillings, M. H. Feinstein, J. L. Garfield, E. L. Rissland, D. A. Rosenbaum, S. E. Weisler and L. Baker Ward. MIT Press: 1987. Pp. 533. \$25, £22.50.

THE enterprise known as cognitive science is constituted out of five contributory disciplines: psychology, linguistics, computer science, philosophy and neuroscience. Given that the last named of these is itself compounded of a variety of established specialisms, the complete cognitive scientist has to encompass a lot of material. Aspirants to that status can, therefore, look warmly upon this, the first book to provide an introduction to all of the relevant disciplines within a single set of covers. ►

The first chapter offers a statement of certain assumptions which together define the enterprise itself. Cognitive scientists believe that the mind can be usefully conceived of as a system for processing information, and that meaningful behaviour can result from formal operations on symbol structures that are in some way representational of the world. The case for such assumptions is explained in the first instance with reference to the use of algorithms for long multiplication, and also by means of the computer metaphor of mind. However, it is stressed right from the start that current hardware and software should not be expected to provide good models of the cognitive architecture, the form of which, it is believed, will be revealed only by way of empirical methods.

The main body of the book consists of 11 individually authored chapters offering tutorial reviews of particular areas. Cognitive psychology, linguistics and computer science are emphasized, with single chapters allocated to philosophy and neuroscience. Inevitably, given the large tracts each chapter must cover, the density of information is high. *Cognitive Science* is not, and could not be, an easy read; there are just too many complex ideas to be put across. Nevertheless, and despite variations in individual style, the quality of exposition is also high throughout, such consistency surely owing something to the editorial hand of the senior author. His opening chapters on cognitive psychology establish the standard, and his concluding chapter on vision gives an unusually accessible account of the how and why of such complex matters as the Laplacian operator and the computing of convolutions.

A further attractive feature of the book, and one that may help the flagging undergraduate reader to keep going, is the invigorating air of optimism with which it is imbued. The authors write with a conviction that progress in cognitive science is real and cumulative. Coupled with this enthusiasm, however, is a tendency to gloss over problems and limitations. In the chapter on knowledge representation, for instance, the legal advice program HYPO is cited as an exemplar of artificial intelligence. A section of dialogue concludes with HYPO outputting two wholly trivial pieces of advice, but no comment is passed either on this particular case or on the extent to which trivial conclusions are or are not typical of supposedly intelligent programs.

There is a wealth of high-grade material to be extracted from *Cognitive Science*. But readers will sometimes have to exercise their own judgement as to what is genuine ore and what fool's gold. □

Art of the soluble

Russell Moseley

The Scientific Attitude. By Frederick Grinnell. Westview, Boulder, Colorado: 1987. Pp. 141. Hbk \$29.95, £26; pbk \$13.50.

TOWARDS the end of his short book, Frederick Grinnell touches upon the inadequacies of a traditional education in science in conveying a true picture of what actually constitutes scientific practice. The image of science, as he points out, is typically idealized while a Whiggish view of progress is attributed to something called 'the scientific method'. Little emphasis is placed on how past progress was accomplished, and even less on why developments took the direction they did. Professor Grinnell hopes that his book will give undergraduate students a better picture of how science is done while, at the same time, providing a guide through the complexities of a scientific apprenticeship for graduate students, and also encouraging practising scientists to reflect on the conduct and direction of modern science.

He approaches this formidable task by looking at scientific activity at the levels of the individual, the collective and society. His account of the third of these is the least convincing. A token gesture towards the interaction of science and politics (including the obligatory reference to the Lysenko affair), a brief look at science, religion and ethical issues, and a curious section entitled 'The Impact of the World on Science' (confined to a single page), do

not take us very far. This is a pity, since the earlier descriptions of practising science as an individual and as a member of a scientific collective have much to commend them.

Certainly, anyone contemplating a career in science could read with profit the account of how scientific collectives operate, although I suspect that the description of graduate programmes, thesis advisers, the difficulties of publishing and the awarding of grants might be as familiar to professionals in the social sciences and humanities as to their scientific colleagues.

The earlier parts of the book will, therefore, be useful in giving students (especially in the life sciences, from which almost all the examples are drawn) a feeling for what actually doing science might entail. It has the merit of being jargon-free and clearly written by someone who, as a professional scientist, will have a credibility with science students that a philosopher or sociologist might not. That the book pays little attention to the interaction between science and its broader social and cultural context can always be remedied by encouraging readers to turn to additional sources. Indeed, many will already be familiar with C.H. Waddington's book which was published almost fifty years ago under the same title as Professor Grinnell's. After reading both books, students might well be encouraged to reflect on the ways in which the scientific attitude has changed over the past half century. □

Russell Moseley is Registrar at the Council for National Academic Awards, 344-354 Gray's Inn Road, London WC1 8BP, UK.

A. R. H. Gellatly is a Lecturer in the Department of Psychology, University of Keele, Staffordshire ST5 5BG, UK.

European heavens

Robert Connon Smith

Fundamental Astronomy. Edited by H. Karttunen, P. Kröger, H. Oja, M. Poutanen and K.J. Donner. *Springer-Verlag*: 1987. Pp. 478. DM 128, £47, \$45.

ONE OF the important differences between North American and European universities is that in the United States and Canada the great majority of students studying astronomy are liberal arts majors with little or no command of mathematics. The descriptive nature of textbooks for these courses is well known; there are many excellent examples, but they are not very suitable for the typical European student of astronomy, who has a training in basic mathematics and physics. For such students, there is a variety of advanced texts but a dearth of elementary accounts that include mathematics.

Fundamental Astronomy goes a long way towards filling that gap. The topics covered are fairly standard for an introductory text, ranging from spherical astronomy to cosmology, via instruments, planets, stars and galaxies. However, the balance is better than usual. Instead of the stress on the Solar System that is common in descriptive texts, space is given to more mathematical topics, such as radiation mechanisms and celestial mechanics, with a brief but useful chapter explaining the different ways of defining temperature.

Unusually, the book has a large number of authors. It evolved from lecture notes

by one of the editors, but has grown with the help of no fewer than 13 colleagues other than the five editors. This wide range of expertise gives the book an authority that would be almost impossible for a single-author text. The translation from the original Finnish is excellent.

Although the treatment makes integral use of mathematics, much of the book would be accessible to someone who has followed that subject to an intermediate level at secondary school. As a result, *Fundamental Astronomy* would also be a good choice for the serious amateur who wants more than just description but whose mathematics cannot cope with more advanced texts. As the editors point out, the advent of personal computers has led to a need for amateurs to have access to exact, but comprehensible, mathematical formalism for their programs.

There is a series of helpful appendices, ranging from elementary calculus to the Lorentz transformation, and including 24 tables of basic information. Some of these are the standard numerical values of physical constants (in SI units); others are more unusual, such as the lists of large telescopes (optical, radio and millimetre). In the text itself, there are other aids to the reader: worked examples are provided at the end of more than half the chapters, while starred sections in small print take the inquisitive reader beyond the general level of the book. Finally, the volume is enhanced by its illustrations, in particular by the central colour supplement. □

Robert Connon Smith is a Lecturer in the Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK.

Stars in a jar

Grenville Turner

Chemistry of the Solar System: An Elementary Introduction to Cosmochemistry. By Hans E. Suess. *Wiley*: 1987. Pp. 143. £21.50, \$24.95.

COSMOCHEMISTRY is the subject that brings astronomy and astrophysics into the laboratory. By studying the chemical, isotopic and mineralogical make-up of primitive meteorites, cosmochemists have been able to provide evidence on which to base theories of the formation of the elements in stellar interiors, as well as theories of the formation of the Solar System from the debris of these stellar furnaces. Professor Hans Suess has been a major contributor to the subject for over half a century, so the publication of an introductory text by him is a significant event. The book is based on a course of lectures given from 1968 to 1985 at the University of California, San Diego.

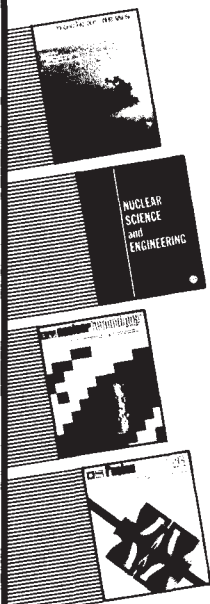
The text is divided into two parts. The

first covers the nuclear physics related to elemental and isotope abundances and element synthesis, while the second concentrates on various aspects of the chemistry of the early Solar System, deduced principally from meteorite studies, though there is also a brief discussion of the planets, asteroids and comets. The content is somewhat unusual and clearly reflects an unashamed bias towards those aspects on which the author has worked. Professor Suess has had a hand in so many areas that this approach gives a reasonable coverage, although the balance could have been better, and as the end of the book approaches subjects are covered with increasing speed and decreasing depth.

The bulk of the first part is given over to a detailed discussion of nuclear stability, shell structure and element abundance systematics, all areas in which Professor Suess was a leading figure in the 1950s. The treatment of mechanisms of nucleosynthesis is in contrast perfunctory, and is largely restricted to an elementary discussion of the r- and s-processes. In view of the current flood of evidence for pre-solar isotopic anomalies in meteorites, assess-

ADVERTISEMENT

The ANS Journals



Your number one source for current information about

NUCLEAR ENERGY

NUCLEAR NEWS
NUCLEAR SCIENCE & ENGINEERING
NUCLEAR TECHNOLOGY TRANSACTIONS
REMOTE SYSTEMS TECHNOLOGY PROCEEDINGS
FUSION TECHNOLOGY

For free catalog contact:

AMERICAN NUCLEAR SOCIETY
 555 N. Kensington Avenue
 La Grange Park, Illinois 60525 USA
 312-352-6611 Telex: 4972673

Reader Service No.30

ment of the implications for other nucleosynthetic processes is noticeably absent. The account of isotope anomalies is itself very patchy with several pages given over to the noble gases and virtually nothing to the so-called FUN anomalies in elements such as calcium and titanium.

The weakness of the book is that it ignores some of the more recent advances in the subject. Some of these omissions, while arguably minor in the overall context, are particularly irritating. The discussion of sources of ^3He in the Earth's atmosphere, for example, ignores the observation made in the author's own institute that (primordial) ^3He is being released into the oceans by volcanic activity at the mid-ocean ridges.

The omission of recent work is perhaps not altogether a bad feature, in that the real strength of book is the way it provides a perspective on cosmochemistry covering the past 50 years (half of the references are to papers published before 1954). In case my review seems too negative, let me conclude by saying that the book is well written and, apart from a few minor errors, enjoyable to read. Combined with one of the modern texts on meteorites, it is a good introduction to an exciting and rapidly moving subject. □

Grenville Turner is a Professor in the Department of Physics, University of Sheffield, Sheffield S3 7RH, UK.

Greek independence

Michael Rowan-Robinson

Cosmology: The Structure and Evolution of the Universe. By G. Contopoulos and D. Kotsakis. Translated by M. Petrou and P. L. Palmer. Springer-Verlag: 1987. Pp. 235. Pbk DM58, £21, \$32.50.

Discovering the Universe. By William J. Kaufmann III. W. H. Freeman: 1987. Pp. 385. Pbk. \$24.95, £24.95.

COSMOLOGY has changed dramatically in the past decade. The advent of Grand Unified Theories, inflation, axions, photinos and other -inos, strings and superstrings, has turned the subject from an impressive synthesis of general relativistic models of the Universe with observational evidence for a hot Big Bang, into a hot-bed of wild speculation about the particle physics of the very early Universe.

As it will take some time for these ideas to settle down, and probably even longer to find any observational evidence for them, this is a difficult moment to be bringing out a new book on cosmology. Contopoulos and Kotsakis's *Cosmology* is a translation of a revised and updated version of their 1982 book in Greek. It survives the pitfalls of this fashion-conscious era by pursuing a rigorously independent-minded attitude to contemporary ideas. The level is introductory undergraduate, with rather little mathematics but a strong physical basis.

The authors have divided their material into three parts, "Observations", "Theory" and "Fundamental Problems". The first part is adequate, though somewhat derivative of other texts. It has not been updated as well as the rest of the book and is especially bad on the late stages of stellar evolution. Stars are said to evolve to black holes if they are more than three solar masses: planetary nebulae are created "by an explosive release of matter from the central star".

The second part, on general relativity and cosmology, provides a good account of modern theoretical ideas, from rotating black holes to Grand Unified Theories and inflation. The third part is an excellent and profound discussion of the fundamental problems of cosmology: the universality of the laws of nature, the uncertainty principle, causality, the origin of time and time's arrow, teleology, the anthropic principle, metaphysics. For anyone with a philosophical turn of mind this section alone makes the book essential reading. And it gives me great pleasure to come across scientists who can write, in discussing Socratic scepticism:

the basic problems of man are the meaning of life and death. Such problems can not be solved either with unjustified optimism or with an

arrogant approach. The inadequacy of naive agnosticism becomes clear in matters like this. Such problems give us a glimpse of new aspects of reality that are not less important than the laws of physics.

Congratulations are due not only to the authors, but also to the translators, Maria Petrou and Phil Palmer, for bringing us an English edition of this original book.

Discovering the Universe by William Kaufmann is a condensation of his earlier *Universe* (which, as it happens, has just appeared in a second edition), to provide a one-term introductory course on astron-

omy for non-scientists. One has to admire the professionalism of the illustrations, which have been brought up to date to include some of the IRAS images. However the book lacks both the soul and the accuracy of *Cosmology*. Life is too short to spend much time with someone who writes of black holes: "Gravity around one of these massive stellar corpses is so strong that it punches a hole in the fabric of space". □

Michael Rowan-Robinson is a Professor in the Department of Mathematical Sciences, Queen Mary College, Mile End Road, London E1 4NS, UK.

Distant demands

Paul M. Mather

Introduction to the Physics and Techniques of Remote Sensing. By Charles Elachi. Wiley:1987. Pp.413. \$42.50, £46.95.

Satellite Remote Sensing: An Introduction. By Ray Harris. Routledge & Kegan Paul:1987. Pp.220. Hbk £22.50, \$62.50; pbk £10.95, \$27.50

REMOTE sensing draws together specialists from many fields. Physicists provide the theoretical basis for the interpretation of data derived from sensors carried by aircraft and satellites, while electrical engineers are concerned with the practical aspects of instrument manufacture. Oceanographers, geologists, agronomists, ecologists, meteorologists and climatologists all interpret and use remotely sensed data, while mathematicians, statisticians and computer scientists are involved with techniques of data processing.

The 'remote sensing specialist' is thus a mythical creature; the discipline is essentially collaborative rather than exclusive. Just as remote-sensing techniques employ electromagnetic wavelengths ranging from centimetres to nanometres, so specialists in the subject range across a broad spectrum of interests and abilities. The two books under review address different ends of this spectrum.

Elachi envisages that his readers will have attained a junior level in physics, specifically that they will have taken introductory courses on electromagnetic and quantum theory. His book aims to provide the scientific and engineering background for students and researchers interested in the basic physics of wave-matter interactions, data collection techniques and applications of remote sensing. Coverage of physical principles and of sensor technology is excellent, although some examples are limited and there is no discussion of data-processing techniques.

In the first part of the book, Elachi

covers basic physical principles, followed by studies of the use of remote sensing in solid- (including ocean-) surface studies. Five chapters in the second part are devoted to atmospheric and ionospheric sensing, using examples from several planets. The quality of the black-and-white illustrations is surprisingly low, although in partial compensation there is a 16-page colour section. In the book of this nature, which is likely to be widely used for reference, a comprehensive index is essential; unfortunately, that provided by Elachi runs to only four pages, which is quite inadequate.

Harris's book is aimed not at the physical scientist or engineer but at undergraduate-level environmental scientists, particularly geographers, geologists and climatologists. The first part contains a sound coverage of physical principles, sensors and platforms, while the second is devoted to a number of examples of applications drawn from the literature.

The level of presentation is considerably lower than Elachi's and reflects Harris's experience as a teacher. His style of writing is clear and enjoyable, and this book should be essential reading for first- and second-year undergraduates who intend to proceed to an advanced remote-sensing course. Like Elachi, Harris provides only a minimal index, and although the quality of the illustrations is good, the benefit of the two-page colour section is dubious.

Neither author considers remote sensing within the broader context of geographical information systems, although it is clear that the synthesis of digital spatial datasets using advanced information-handling techniques will be the theme for the 1990s. But in terms of their own specific aims, both are successful. Elachi has provided a basic reference volume which will be perused, rather than read from cover to cover, by researchers. Harris's book, on the other hand, is a gentle introduction to an exciting, fast-changing and essentially multidisciplinary field. □

Paul M. Mather is a Senior Lecturer in the Department of Geography, University of Nottingham, Nottingham NG7 2RD, UK.

Mixed reactions

Malcolm Scott

Nuclear Physics for Engineers and Scientists: Low Energy Theory with Applications Including Reactors and Their Environmental Impact. By S. E. Hunt. *Ellis Horwood*: 1987. Pp.371. £51.50, \$116.40.

Principles of Nuclear Science and Engineering. By A. A. Harms. *Research Studies Press, Letchworth, Herts/Wiley, New York*: 1987. Pp.192. £22, \$43.05.

ALTHOUGH the basic aim of these two books is similar, namely provision of the physics background necessary to understand fission and fusion reactors, the approaches and achievements could not be more disparate. Many scientific books claim to be written simply, with little or no assumed mathematics, whereupon the author starts by solving a second-order differential equation. Professor Hunt states in his short introduction that he wishes to emphasize the physical interpretation of nuclear behaviour, in order to make the subject intelligible to a non-physicist — and he proceeds to do just that. Indeed, the text has an elegance, yet simplicity, of presentation which one must admire and would have to strive very hard to emulate. Not only does Hunt explain the underlying principles clearly, but he manages to do so in a very readable way, introducing small personal touches which help to take away some of the impersonality associated with science (“... many man years, including some of the author’s own, were expended on ...”).

The book starts with a chapter on the nuclear atom, and the next six are devoted to basic nuclear and atomic physics. There follow five chapters dealing with the production, properties and detection of charged particles, photons and neutrons. The physics of fission and fusion is covered in two separate chapters, while the concluding three deal with commercial fission reactors, fuel reserves (nuclear and non-nuclear) and nuclear safety.

In a book of this breadth, an author cannot have had first-hand professional involvement in every topic discussed. There are, however, only a few instances where this has led to statements which are slightly questionable, or where the emphasis gives a misleading impression. There are, surprisingly, a number of typographical errors, and a whole chapter has the wrong running headline. Most of the slips are easily spotted, but if the mean current of the Nuffield Cyclotron were really 500 MA we would be running a power station not a research accelerator!

Professor Hunt’s book is valuable for several different classes of reader. It could serve both as refresher course for those who have already studied the field, or as

an introduction for those who have not. Its breadth and lucidity are impressive, and it achieves its depth of understanding without resorting to recipes. Only the cost of the hardbound edition is off-putting. One can only hope that a paperback version will soon be forthcoming, so that the book can reach the readership it deserves.

In contrast, Professor Harms’s book starts with a relatively long preface in which he states his pedagogical intention of applying a “bottom-up” approach to nuclear fission and fusion, contrasted with a “top-down” approach, involving “simplification of a complex and specialised subject by a variety of conceptual reductions and analytical approximations”.

With such a claim, I expected a text either having simple rigour or offering insights disguised by alternative approaches. Unfortunately, this is not the case. One reason is simply the prose style, which is often long-winded and ponderous. Another is that, unlike in many modern textbooks, areas which may be conceptually difficult have not been presented with the potential problems

faced by the reader clearly in mind. For example, in discussing the question of fission reactor criticality the same symbol is used for materials and geometric buckling, thus obscuring the distinction between them, a confusion only partly resolved by the text. Only in one of the problems at the end of the chapter is the origin of this distinction hinted at, and the reader is left to puzzle things out. Similarly, the reason for the fundamental mode describing the spatial distribution of the neutron flux in a critical assembly is only alluded to, yet giving the time-dependent solution would both make the reason clear and, I would have thought, represent a true “bottom-up” approach.

Unfortunately, one can cite many instances where simple points are either obscured or are laboured. Given the range of the textbooks which already cover this field, it is, sadly, difficult to recommend this one. □

Malcolm Scott is a Senior Lecturer in the Department of Physics, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

Odd, isn’t it?

Tony Sudbery

From Paradox To Reality: Our Basic Concepts of the Physical World. By Fritz Rohrlich. *Cambridge University Press*: 1987. Pp.227. £25, \$34.50.

Particles and Paradoxes: The Limits of Quantum Logic. By Peter Gibbins. *Cambridge University Press*: 1987. Pp.181. Hbk £25, \$34.50; pbk £8.95, \$11.95.

SINCE the early decades of this century, the popular image of theoretical physics has been surrounded by an aura of paradox. The two revolutionary theories of that period, relativity and quantum mechanics, seemed to violate not only common sense but all rationality. In consequence the basic concepts of theoretical physics are an appropriate subject of serious study for more than just physics students, and these two books offer reading matter for different parts of this wider audience.

Fritz Rohrlich’s book is based on an undergraduate course “Concepts in Contemporary Physics” that he offers to students majoring in non-science subjects. As well as teaching them about the specific concepts of relativity and quantum mechanics, he is concerned to give a general understanding of the way scientific theories fit together and what happens when a theory is superseded in a scientific revolution. Rohrlich’s view is that this is bound to happen to every theory, and that only when it does is the theory complete, or ‘established’, because only then does it acquire a definition of its domain of valid-



Erwin Schrödinger, one of the Men Who Made a New Physics, featured in a book recently reissued in paperback by University of Chicago Press, price £9.50, \$11.95.

ity. This is a clear and coherent account of a doctrine which is often impressed on science students, though I hope some of Rohrlich’s readers will wonder if it really is the necessary truth he says it is.

In his accounts of special relativity and quantum mechanics, Rohrlich focuses on the central concepts. He places both theories in their historical context, describing their roots in earlier ideas and the subsequent delimitation of their validity by general relativity and quantum field theory, respectively. The author does not assume any knowledge of physics, but his readers are expected to work (and they may be puzzled by the physicist’s use of such words as ‘interfere’ and ‘scatter’).

The paradoxical aspects of the concepts are confronted and, in the case of special

relativity, resolved in an excellent discussion of the twin paradox. In quantum mechanics Rohrlich deals with Schrödinger's cat paradox and the Einstein-Podolsky-Rosen experiment, and confidently identifies the points at which these arguments 'fail'. Not all his scientific readers will agree with him, and his non-scientific readers should perhaps have been told that in this area there is no general consensus among physicists.

Rohrlich devotes a section to quantum logic, the notion that the empirical success of quantum mechanics forces us to abandon classical logic in the same way that general relativity forces us to abandon Euclidean geometry. Quantum logic is the subject of the second half of Peter Gibbins's *Particles and Paradoxes*; the first half is an account of quantum mechanics for philosophy students. Gibbins makes some interesting points about the place of the philosophy of quantum mechanics, and his development of quantum logic will be of interest to readers with some knowledge of formal logic. Unfortunately, his account of quantum mechanics contains mistakes and confusions, and is written in a disjointed style which is hard to follow in places. Nevertheless, as a book for philosophy students on this most philosophically puzzling part of physics it can only be welcomed. □

Tony Sudbery is a Lecturer in the Department of Mathematics, University of York, Heslington, York YO1 5DD, UK.

In the beginning

John Tarney

The Young Earth: An Introduction to Archaean Geology. By E.G. Nisbet. Allen & Unwin: 1987. Pp. 402. Hbk £40, \$60; pbk £17.95, \$34.95.

The Young Earth is really about old rocks — those formed during the Archaean period, more than 2,500 million years ago. Whereas geologists feel they are beginning to have a reasonable understanding of how the modern Earth works in this era of plate tectonics, the picture becomes increasingly hazy as we go back in time. It is not always easy to interpret the rocks in terms of modern tectonic processes, and more comprehensive observations are required which have to be integrated with existing knowledge to formulate new or modified processes. Thus models of the early Earth are necessarily tempered with speculation, and the author freely admits to the difficulty by describing his book as "a personal view".

Nonetheless, this is one of the most comprehensive texts ever produced on the Archaean. Nisbet appears to explore every possible avenue to uncover the

Crystal clear

R.A. Howie

Optical Mineralogy, Vols 1 and 2. By Ernest G. Ehlers. Blackwell Scientific: 1987. Vol. 1 Theory and Techniques, pp. 158, \$39.95, £32. Vol. 2 Mineral Descriptions, pp. 286, \$49.95, £38.

Optical Mineralogy provides in two volumes the necessary information to identify the commoner rock-forming minerals, both in thin section and grain mounts, by use of the petrographic microscope. Both volumes are illustrated by numerous line drawings, and both have the same colour plates of the Michel-Lévy chart of polarization colours and of interference figures, with the addition in Vol. 2 of 70 colour photomicrographs (these are of high quality but remote from the text descriptions of the minerals shown). The work is thus similar in scope to Shelley's *Optical Mineralogy* (Elsevier, 1985) and Nesse's *Introduction to Optical Mineralogy* (Oxford University Press, 1986).

In Vol. 1, after introductory chapters on the polarizing microscope and on the nature and behaviour of light, there are sections on the examination of isotropic substances, uniaxial and biaxial minerals, the relation between the optic behaviour of biaxial materials and crystal symmetry, and the microscopic identification of unknown minerals. This volume ends with a chapter dealing with the spindle stage and (briefly) with the universal stage. Readers get the benefit of a text written by an experienced teacher who fully appreciates the problems likely to be encountered by students, particularly when dealing with the biaxial indicatrix.

facts, assess their significance, link them together and weave them into plausible models. He describes typical Archaean low-grade greenstone belts and high-grade granite gneiss terrains, supplying some of the detail from his own observations. He then marshals together field, geochronological, isotopic and trace-element evidence from Archaean volcanic, plutonic and sedimentary rock associations to develop ideas about the nature and evolution of the Earth's mantle and the growth and development of the continental crust. A whole chapter is devoted to Archaean mineral deposits (gold, copper, nickel and so on) in recognition of their importance as a world resource.

A large part of the book, however, is devoted to the evidence for life in these ancient rocks. This theme is later developed into a discussion of molecular palaeontology and a review of various ideas on the origin of life, and the influ-

ence living organisms have had on the development of the planet itself — particularly the atmosphere and hydrosphere. It is useful to have this essay tied in with more fundamental aspects of the growth and development of the Earth's crust and early crustal processes. The result is a comprehensive view of the early stages of evolution of the planet.

To make the book suitable for the general reader, passages of detailed scientific argument are interlaced with sections of chatty text and quotations from classical, biblical or even more unusual sources. A glossary of geological terms is provided for those who might get bogged down with the jargon. Popularization is, however, not excessive: it is not Page Three but the very last page which tells us that sex began in the Proterozoic. □

The books are well set out, with numerous black-and-white photographs as well as sketches. Together the two volumes are expensive, but a student might well consider settling for Vol. 2 and hope to find Vol. 1 in the library. □

R.A. Howie, Department of Geology, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, UK, is Emeritus Professor of Mineralogy in the University of London.

ence living organisms have had on the development of the planet itself — particularly the atmosphere and hydrosphere. It is useful to have this essay tied in with more fundamental aspects of the growth and development of the Earth's crust and early crustal processes. The result is a comprehensive view of the early stages of evolution of the planet.

To make the book suitable for the general reader, passages of detailed scientific argument are interlaced with sections of chatty text and quotations from classical, biblical or even more unusual sources. A glossary of geological terms is provided for those who might get bogged down with the jargon. Popularization is, however, not excessive: it is not Page Three but the very last page which tells us that sex began in the Proterozoic. □

John Tarney is F.W. Bennett Professor of Geology at the University of Leicester, Leicester LE1 7RH, UK.